DOCUMENT RESUME

ED 187 157 HE 012 206

AUTHOR	Newble, David I.
TITLE	Evaluation of Clinical Competence.
INSTITUTION	Adelaide Univ. (Australia).
PUB DATE	77
NOTE	79p.
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Behavioral Objectives; *Clinical Experience; *Competence; Competency Based Education; Criterion Referenced Tests; Definitions; Educational Objectives; Foreign Countries; Higher Education; *Medical Education; *Medical Students; Models; *Physician Patient Relationship; Student Certification; *Student Evaluation
IDENTIFIERS	*Australia; University of Manitoba (Canada)

ABSTRACT

In Australia the usual evaluation of whether a student will perform adequately as a doctor is a subjective evaluation of his clinical performance, usually at the completion of five or six years at medical school. The evaluation is performed on an inadequate and uncontrolled patient sample and appears to be subject to many errors. Recent work investigating and attempting to define "clinical competence" has produced alternative ways of approaching the problem of evaluation. At the same time, new methods of testing competence have been developed and are being integrated into various certifying examinations. This report reviews the current situation in evaluating and defining clinical competence, and describes attempts to develop a model suitable for producing a certifying examination for Australia. It is hoped that the model may have a wider application for graduate trainees and practicing physicians. Appended are outlines of: categories of clinical performance; elements of the physician patient relationship; criteria for classifying physician actions; a definition of clinical competence; a statement of behavioral objectives for evaluation; terminal learning objectives of the Department of Medicine of the University of Manitoba; a "problem blueprint" for examination purposes; and test selection in the assessment of clinical competence. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

EVALUATION OF CLINICAL COMPETENCE

1. A review.
2. The development of a model for a problem based, criterion referenced examination for evaluating the clinical competence of medical students.

David I. Newble

ADVISORY CENTRE FOR UNIVERSITY EDUCATION
THE UNIVERSITY OF ADELAIDE

© 1977, ACUE, The University of Adelaide

Reprinted 1979

All rights reserved – no part of this publication
may be reproduced or transmitted, in any form or
by any means, without permission of the Copyright Owner.

ISBN 0 909688 19 2

Typesetting and printing by
The University of Adelaide,
Information Office.

ACKNOWLEDGEMENTS

I am indebted to the many people concerned with medical education who gave me advice, information and experience during my study leave in Britain, the U.S.A. and Canada. I would particularly like to thank Professor Stephen Abrahamson and his colleagues of the Division of Research in Medical Education at the University of Southern California, Dr. Vic Neufeld and other members of the Program for Educational Development at McMaster University, Faculty of Health Sciences and Professor Lee Shulman and other staff members of the Office of Medical Education, Research and Development at Michigan State University. I would also like to thank Dr. John Lloyd of the Division of Postgraduate Medicine at the University of Southern California.

The period of study was supported by the University of Adelaide Study Leave Fund and I am grateful to the University for allowing me to take an additional three months leave and to accept an Overseas Research Fellowship awarded by the Education Research and Development Committee. This award has allowed me to travel to many centres conducting medical educational research and has contributed greatly to my knowledge and experience.

Finally I would like to thank Dr. William Hall of the Advisory Centre for University Education at the University of Adelaide for his help in preparing the manuscript for publication.

This book is based on a report prepared during a period of study leave from the Department of Medicine at the University of Adelaide in 1975 and while in receipt of an Overseas Fellowship awarded by the Education, Research and Development Committee.

CONTENTS

		Page
Chapter 1	Introduction	7
Chapter 2	Towards a Definition of Clinical Competence	8
	References	12
Chapter 3	Methods Used in Evaluating Clinical Competence	14
	Introduction	14
	Test Characteristics	15
	Tests for Assessing Clinical Competence	15
	1. Essay	16
	2. Completion or Short Answer	16
	3. Objective (Multiple Choice)	16
	4. Paper Simulation	17
	5. Computer Simulations	18
	6. Other Simulations	20
	Simulated patients	21
	Simulation devices	22
	Patient games	24
	7. Oral Examinations	25
	8. Direct Observation	26
	9. Record Audit	27
	10. Multiple Methods	28
	11. Outcome Measures	29
	References	30
Chapter 4	Examinations of Clinical Competence	39
	References	42
Chapter 5	Model for Developing a Problem-Based Criterion Referenced Test of the Clinical Competence of Medical Students	44
	References	53
	Appendix 1	54
	Appendix 2	56
	Appendix 3	59
	Appenidx 4	63
	Appendix 5	65
	Appendix 6	69
	Appendix 7	75
	Appendix 8	77

1. INTRODUCTION

In Australia and in many other countries, the usual way of determining whether a student will perform adequately as a doctor is based on a subjective evaluation of his clinical performance, usually at the completion of five or six years at the University Medical School. The evaluation is performed on an inadequate and uncontrolled patient sample and there is ample evidence that this evaluation, the traditional medical viva, is subject to many errors. In some schools the medical viva has been partly or fully replaced by in-course assessment during several clinical clerkships but the reliability of such evaluation is equally open to question. In most schools the clinical evaluation is complemented by an objective (multiple choice or equivalent) examination predominantly testing the recall and recognition of factual knowledge.

Over the past few years, many individuals and organisations have become increasingly aware of the deficiencies of the evaluation methods used in evaluating clinical performance, not only at the undergraduate level, but also at the level of postgraduate training. The increasing demands of the community and government for accountability of all doctors makes the development of reliable and valid measures of competence even more important.

It is the obligation of Medical Faculties in Australia to assure the State Registration Boards that each students has attained a certain minimum level of clinical competence. In the United Kingdom a similar situation applies. In the U.S.A. and Canada, however, all students are required to pass the examinations set by the National Board of Medical Examiners or the Medical Council of Canada. The emphasis in recent years has been on developing, and in most cases achieving, reliable and objective examinations but this has often been at the expense of validity.

Recent work investigating and attempting to define "clinical competence" has produced alternative ways of approaching the problem of evaluation. At the same time, new methods of testing competence have been developed and are being integrated into various certifying examinations. This report will attempt to review the current situation in the areas of defining and evaluating clinical competence and then attempt to develop a model which would be suitable for producing a certifying examination appropriate to the conditions and needs of a medical school in Australia. It is hoped that the model may have a wider application in developing test formats for evaluating clinical performance in postgraduate trainees and practising doctors.

2. TOWARDS A DEFINITION OF CLINICAL COMPETENCE

It is impossible to arrive at a simple or single definition of clinical competence as it will vary depending on the different requirements of the individual or organisation concerned in making the definition. For instance, the definition applicable to a medical student will be different to that of a practising specialist. In recent years there has been less of an attempt to define clinical competence in the restricted sense of the word – a statement of the precise meaning of the term – but more of an attempt to identify the components of clinical competence. There have also been significant changes in the philosophy and attitude of medical educators concerning what should be included among these components of clinical competence.

Two basic approaches have emerged. The first, and most extensively investigated, is one of identifying and defining the components of the process which leads to satisfactory clinical performance. The second is one which is concerned with the outcome of clinical performance. It is obvious, however, that both must be included in any comprehensive statement about clinical performance (Barro, 1973).

For the undergraduate student, there have been few explicit definitions or statements about what he is expected to perform and at what level of proficiency. He has traditionally relied on the "definitions" implicit in the format and content of the examinations or he has relied on the apparent expectations of his clinical teachers. The teachers, often unwittingly assuming this role, have usually failed to define precisely what they intend the student to know, what they expect him to be able to do and what attitudes they expect him to develop. In fact, the evidence to date has shown that, at least in Britain, there is an inverse relationship between what is taught and what is used in clinical practice (Hodgkin, 1966). In North America, grades at the university have been shown to have little or no correlation with future clinical performance (Price et al., 1971; Peterson, *et al.*, 1956).

In an attempt to overcome such difficulties, educators have stressed the importance of defining objectives and defining these in terms which describe what the student should be able to do at the conclusion of his training. These are known as behavioural or terminal objectives (Mager, 1962; 1967). They are usally expressed as a series of statements which commence with a phrase like: "The graduating students will be able to ..." and conclude with a specific skill and expected level of performance. Despite the logic of such an approach and the monumental efforts of some schools to define such objectives (e.g. Abraham

Lincoln School of Medicine, University of Illinois) there appears to have been difficulty in implementing this approach at the instructional and evaluation level.

A different approach has involved various methods of observing, recording and classifying the actual performance of a variety of doctors as they deal with patients and their problems. One such method is that of task analysis. Examples include studies where doctors kept log diaries of their own activities (Nelson, 1973; Adams and Mendenhall, 1974) or where information was recorded by trained observers (e.g. Peterson, 1956). Studies like these produce lists of problems or conditions with which the doctor has to deal and skills which he uses in order to cope with them. The concern of most of this work has been to identify practice profiles (Sivertson, 1973) and to develop criteria of good medical practice (Peterson, 1956). No task analysis studies have been identified in the literature relating to *the work of the intern (junior resident) which is the job the student will be expected to fulfil when he graduates.* Kane et al. (1973) performed a task analysis of family practitioners to help in the development of learning objectives and to support a concept of producing "comprehensive care physicians". Certainly, it should be possible to make generalisations applicable to medical students from studies of doctors in independent practice but most graduating students will work for at least one year in supervised hospital practice as interns and it is a moot point whether the University should be training the student for this role or for a wider role as a "comprehensive care physician".

The technique which has been most successful in establishing classifications (taxonomies) of clinical competence has been the critical incident technique developed initially by Flanagan (1954) and used extensively by Hubbard et al. (1965), Sanazaro & Williamson (1968,1970) and Sanazaro & Bates (1968). The method consists of identifying incidents of clinical performance that are conspicously good or effective and that are conspicuously bad or ineffective. The technique has some limitations because the incidents are not necessarily a representative sample of actual experience, but clearly the incidents do not include irrelevant categories and are usually clearly related to outcomes of patient care. The study reported by Hubbard et al. is the more relevant in the context of undergraduate education. It was conducted by the American Institute of Research for the National Board of Medical Examiners (NBME) and was based on the performance of interns. A total of 3300 incidents were collected from 600 senior consultants and registrars (residents) identifying personally observed good or bad practice of the interns they were supervising. These incidents were then sorted and categorised into nine major areas of clinical performance with additional sub-headings:

1. History
2. Physical Examination
3. Tests and Proeedures
4. Diagnostic Acumen
5. Treatment
6. Judgment and Skill in Implementing Care
7. Continuing Care
8. Physician–Patient Relation
9. Responsibilities as Physician (See Appendix I for details)

These categories have provided a framework for subsequent developments and changes in the clinical component (Part 3) of the National Board examination.

The study by Sanazaro and Bates was performed on students undertaking a comprehensive medicine course and produced a classification into four main categories and many detailed sub-categories. The main categories were:

1. Collecting and Interpreting Clinical Information
2. Instituting and Maintaining Comprehensive Care
3. Physician–Patient Relationship
4. Responsibilities as a Physician

(See Appendix 2 for details)

Sanazaro and Williamson collected over 9,000 incidents from 2,300 practising specialists in internal medicine, surgery, paediatrics and obstetrics–gynaecology who were all identified as 'good' doctors. The main categories identified were:

1. General Care
2. Evaluation (General)
3. Evaluation (Specific)
4. Treatment (General)
5. Treatment (Specific)

(See Appendix 3 for details)

Detailed inspection of these classifications and sub-classifications reveals a very close similarity of content. Though the categorisations could reflect some bias by the authors they do seem to cover components of performance identified by other techniques (e.g. task analysis; committees of experts).

The information and classifications produced by these critical incident studies appear to provide a practical framework for defining clinical competence in more detail and this will be discussed further in Chapter 5. The classifications have great validity as they are defined from observed clinical outcomes of performance. However, the structure of the classifications may suggest that the process of clinical problem solving and competence occurs in a sequential manner and this concept is apparent in most undergraduate teaching programmes. The traditional approach to student training assumes that certain sequences of teaching and learning are important, e.g. a detailed knowledge of basic medical sciences is a necessary pre-requisite to clinical teaching; clinical interviewing should be taught before the physical examination and both should be mastered before attention is given to teaching diagnostic and treatment skills. Recent work on the problem solving process and innovative teaching programmes based on clinical problem solving such as those seen as McMaster University (Neufeld and Spaulding 1973, Neufeld and Barrows 1974) and Michigan State University (Ways et al. 1973), raise serious doubts about these assumptions and about how we should define, teach and evaluate clinical performance.

The methods of study discussed so far are based on judgments made about a doctor's performance as his behaviour is observed. They do not identify the reasoning processes that lead to this behaviour. In looking at this area, Elstein and his colleagues (Elstein et al., 1972; Shulman and Elstein, 1975) have made contributions which are of particular significance. They have systematically observed expert clinicians undertaking the work-up of a simulated patient. The work-up was videotaped and when completed the tape was replayed to the doctor and he was asked to think aloud and recall his thoughts and reasoning processes. This information was subsequently analysed to produce their theory of medical inquiry. This proposes that experienced doctors generate specific diagnostic hypotheses well before data collection is complete and proceed to test these hypotheses as further information is gathered. Hypotheses were found to be rank ordered on the basis of probability, seriousness, treatability and novelty. The general model encompasses four major activities:

1. Cue acquisition
2. Hypothesis generation
3. Cue interpretation
4. Hypothesis evaluation

Decision making and judgment occur both during and at the conclusion of these activities. They have been able to define similarities with other psychological research on problem solving, judgment and decision making, in particular that of Newell and Simon (1972). Many other factors have been identified which effect human problem solving (Shulman and Elstein, 1975).

Another approach to the study of clinical problem solving has come through the development of paper and pencil patient management problems (Rimoldi, 1963; McCarthy and Gonella, 1967; McGuire and Babbott, 1967). These tests of diagnostic skill have less validity than studies using real or simulated patients but do provide a controlled environment for studying patterns of behaviour during problem solving. One important point that has been established, using such techniques, is that specific knowledge is necessary for effective problem solving and it is therefore unwise to generalise about problem solving ability from single case situations (McGuire and Page, 1973).

In summary, it can be concluded that the process of clinical problem solving and the entity of clinical competence are complex phenomena. It is not possible to define them in simple terms. The development of classifications or taxonomies based on studies of "what doctors do" has value in providing a framework defining the areas in which skill is required. The studies of ways in which doctors develop these skills and use information to solve their patient's problems have less immediate utility in defining clinical competence but have clear lessons for those concerned with the instruction of medical students.

ED187157

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Univ of Adelaide

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

REFERENCES

Adams, F.H., and Mendenhall, R.C., Profile of the Cardiologist: Training and Manpower Requirements for the Specialist in Adult Human Cardiology. Amer. J. Cardiol., 34 389, 1974.

Barro, A.R., Survey and Evaluation of Approaches to Physician Performance Measurement. J. Med. Educ. 48 1050, 1973.

Elstein, A.S., Kagan, N., Shulman, L.S., Jason, H., and Loupe, M.J., Methods and Theory in the Study of Medical Inquiry. J. Med. Educ., 47 85, 1972.

Flanagan, J.C., The Critical Incident Technique. Psychol. Bull., 51 327, 1954.

Hodgkin, K., Towards Earlier Diagnosis. E. and S. Livingstone, 1966.

Hubbard, J.P., Levit, E.J., Schumacher, C.F., and Schnabel, T.G., An Objective Evaluation of Clinical Competence. New Engl. J. Med., 272 1321, 1965.

Kane, R., Ross Woolley, F. and Kane, R., Toward Defining the End Product of Medical Education, J. Med. Educ., 48 615, 1973.

Mager, R.F., Preparing Instructional Objectives. Fearon, 1962.

Mager, R.F., and Beach, K.M., Developing Vocational Instruction. Fearon, 1967.

McCarthy, W.H., and Gonnella, J.S., The Simulated Patient Management Problem, A Technique for Evaluating and Teaching Clinical Competence. Brit. J. Med. Educ., 1 348, 1967.

McGuire, C.H., and Babbott, D., Simulation Technique in the Measurement of Problem Solving Skills. J. Educ. Meas., 4 1, 1967.

McGuire, C.H., and Page, G., An Overview of Physicians Performance on Simulated Clinical Encounters Presented in Patient Management Problem Format. AERA Paper, Feb., 1973.

Nelson, A., Study of the Validity of the Task Inventory Method of Job Analysis. Paper presented to Amer. Public Health Assoc. Meeting, 1973.

Neufeld, V.R., and Spaulding, W.B., Use of Learning Resources at McMaster University. Brit. Med. J. 3 99, 1973.

Neufeld, V.R., and Barrows, H.S., The McMaster Philosophy: An Approach to Medical Education. J. Med. Educ. 49 1040, 1974.

Newell, A., and Simon, H.A., Human Problem Solving. Prentice-Hall, 1972.

Peterson, O.L., Andrews, L.P., Spain, R.S., and Greenberg, B.G., An Analytical Study of N. Carolina General Practice, 1953–54. J. Med. Educ., 31 (Pt. 2), 1956.

Price, P.B., Taylor, C.W., Nelson, D.E., Lewis, E.G., Loughmiller, G.C., Mathiesen, R., Murray, S.L., and Maxwell, J.G., Measurement and Predictors of Physician Performance. LLR Press, 1971.

Rimoldi, H.J.A., Rationale and Applications of a Test of Diagnostic Skills. J. Med. Educ., 1 348, 1963.

Sanazaro, P.J., and Bates, B., Joint Study of Teaching Programs in Comprehensive Medicine. J. Med. Educ., 43 389, 1968.

Sanazaro, P.J., and Williamson, J.W., A Classification of Physician Performance in Internal Medicine. J. Med. Educ., 43 389, 1968.

Sanazaro, P.J., and Williamson, J.W., Physician Performance and Its Effect on Patients. Medical Care. 8 299, 1970.

Shulman, L.S., and Elstein, A.S., Studies of Problem Solving, Judgment and Decision Making. In: Review of Research in Education, Ed. Kerlinger, Publ. Peacock & Co., 1975.

Sivertson, S.E., Meyer, T.C., Hansen, R., and Schoemenberger, A., Individual Physician Profile: Continuing Education Related to Medical Practice. J. Med. Educ., 48 1006, 1973.

Ways, P.O., Loftus, G., and Jones, J.M., Focal Problem Teaching in Medical Education. J. Med. Educ. 48 565, 1973.

3. METHODS USED IN EVALUATING CLINICAL COMPETENCE

Introduction

This chapter will be confined to a review of techniques currently available to evaluate aspects of clinical competence. An attempt will be made to identify the strengths and weaknesses of each method and make some judgments about applicability. In general the discussion will be concerned with measures of the process of clinical performance rather than measurement of the outcome of clinical performance in terms of patient care. It is recognised that the validity of process measures is dependent on their prediction of outcome performance (Williamson 1971, Sanazaro and Williamson 1970). Medical student evaluation has been undertaken entirely on process measures but the few studies that have been performed show little or no correlation between grades and future performance (see review by Wingard and Williamson, 1973). However, it is fair to say that it is difficult to devise outcome measures which could be applied to evaluating medical students as they rarely assume direct responsibility for patient care. What is required are process measures which are valid, i.e. will be predictive of future competence.

Current research is also concerned with developing tests with a higher degree of fidelity – a term introduced into the medical literature by Shulman and his colleagues at Michigan State University. The purpose is to establish examination settings resembling the actual situation in which clinical performance occurs.

The performance of students is sometimes divided into four broad areas (Irby & Morgan, 1974; Barro, 1973) and this type of categorisation is used widely.

1. Knowledge
2. Technical (Psychomotor) skills
3. Interpersonal skills
4. Habits and attitudes

The knowledge or cognitive area has been subclassified by Bloom and his colleagues (1956) into knowledge, comprehension, application, analysis, synthesis and analysis. This taxonomy is of some importance as it has been used freely by objective test item writers as a basis for item classification. Other heirarchies have been devised (e.g. Gagne, 1970) that include the commonly used category of

problem solving. A simplified classification, which appears to have more practical utility, is proposed by Irby & Morgan. They divide the knowledge area into aspects concerned with recall, aspects of interpretation or evaluation of information and into problem solving.

Taxonomies of the psychomotor domain (Harrow, 1972) and the affective domain (Krathwohl et al. 1964) provide detailed subclassifications relevant to the areas of technical and interpersonal skills. These technical and interpersonal skills include all the specific skills which are part of the doctor's job. The extent of these skills are constantly being redefined and this is particularly so at the present time in relation to interpersonal skills and problem solving ability (Feinstein, 1970; Engel, 1971, 1973; Kane et al. 1973; Stead, 1971; Dudley, 1970).

Habits and attitudes are increasingly coming under study and tests are being devised to evaluate these entities.

Test Characteristics

In evaluating the clinical competence of medical students, we are primarily concerned with **achievement** tests which are designed to test present mastery of the subject or of certain skills. They may be used to attempt to predict future performance (certification examinations) or to provide information for feedback purposes (diagnostic examinations). Any test used must satisfy both practical and technical requirements (Adkins, 1974; Hubbard, 1971). The test must be **reliable**. That is, it must produce consistent results. The test must also be **valid**. It must have **content or face validity**, meaning it must be testing the appropriate content or defined objectives and, as mentioned previously, it should predict future performance (**criterion related or predictive validity**). In general, we will be less concerned with the difficult area of **construct validity**. Finally, the test must be **practical** in terms of ease of construction, administration, cost and scoring.

Tests for Assessing Clinical Competence

1. Essay
2. Completion or Short Answer
3. Objective (Multiple Choice)
4. Paper Simulations (Patient Management Problems)
5. Computer Simulations
6. Other Simulations
    a. Simulated Patients
    b. Simulation Devices
    c. Patient Games
7. Oral
8. Direct Observation
9. Record Audit
10. Multiple Methods
11. Outcome Measures

1. Essay

Essay examinations are less commonly used today though they may return, in some form, as interest grows in evaluating higher order learning and attitudes. The main problems of essay examinations are those of adequately sampling the content area, constructing good questions, the time required for marking and, most importantly, the unreliability of scoring. Methods have been devised for improving reliability and include producing a model answer, allocating scores to key points, strategies to avoid the "halo" effect, rescoring random samples and use of two scorers. Unfortunately, the areas where the essay test has particular strength – organisation of ideas, expression of ideas, drawing inferences, creativity – are not areas which lend themselves to objective evaluation (Adkins, 1974). However, if these skills are considered essential and cannot be tested in any other way, essays should probably still be set in order to demonstrate to the student the commitment of Faculty to such skills.

2. Completion or Short Answer

These are essentially an alternative to objective tests requiring the completion of a statement or brief answers to highly specific questions for which an itemised scoring system can be devised (Adkins, 1974). They predominantly test recall. The main advantage is in avoiding the cueing effect associated with objective tests and the main disadvantage is that they require hand scoring, though not necessarily by Faculty staff. We have recently undertaken a comparison of student performance on equivalent short answer and objective tests. Preliminary results suggest that the cueing effect in the objective tests is such as to invalidate their use if one is concerned to measure ability in the clinical area (Newble & Baxter, 1976).

3. Objective (Multiple Choice)

Many good texts and reviews on objective tests are readily available so they will receive only brief consideration here. They are the most widely used of all tests and most faculty members will be familiar with them either as examination participants or from item writing (Adkins, 1974; Hubbard, 1971). The National Board of Medical Examiners (NBME) has probably had the greatest experience with this format and experience has led them to use only three types – one best response, matching response and multiple true/false (Hubbard, 1971). This choice was based on the particular requirements of medical content and the reliance placed on non-expert item writers to devise the questions.

The characteristics of good objective tests are the high reliability of scoring, scoring economy and the ability to screen large content areas. The development of local, national and even international banks of items reduces preparation time. However, there is a growing dissatisfaction with the heavy emphasis on objective tests mainly based on assertions that they measure predominantly recognition or recall of isolated bits of factual knowledge, that they have not been related to meaningful objectives (Levine et al., 1970), and that they do not have predictive or criterion related validity (Wingard & Williamson, 1973).

In an examination where the results of a critical incident study were available to direct item writers on the areas of content and skill that should be tested, a subsequent analysis showed that the majority of questions still measured simple recall. A factor analysis demonstrated that the multiple choice questions only measured one of the five defined aspects of clinical competence. The areas of competence not measured included the affective and skill domains, deductive reasoning and judgment (Levine et al., 1970).

It is concluded the objective tests should usually be used only for sampling the wide area of content knowledge at a recall level. They can be used for higher order knowledge testing if skilled item writers are available. Examples of such test items can be found in Bloom's well known *Taxonomy of Education Objectives* (Bloom et al., 1956). The heavy reliance placed on this test format for certifying examinations should be reduced and other more valid measures developed.

4. Paper Simulations

These tests now provide the most widely used evaluation instrument, other than objective tests, in certifying examinations of clinical competence. The NBME includes two in its Part III examination, and the American Board of Internal Medicine is proposing to use 15 such tests. They also enjoy wide acceptance by students.

With the increasing popularity of such tests has come an increasingly confused terminology. The pioneering work in this field was done by Rimoldi (1963) who developed a *Diagnostic Management Problem (DMP)* during his studies on the problem solving process. The technique was adapted by McGuire and her colleagues (McGuire & Babbott, 1967; McCarthy & Gonnella, 1967) and called a **Patient Management Problem (PMP)**. McGuire and her colleagues have continued to undertake research on this technique and have published a book of examples (McGuire & Solomon, 1971). Many adaptations of these techniques have been reported (Hubbard, 1971; Helfer & Slater, 1971). A recent review identified 18 varieties of PMP (Galofre, 1974).

The typical PMP begins with a variable amount of patient data and identifies the doctor–patient setting. The student is then requested to collect further data in a sequential fashion either in a linear (e.g. Hubbard, 1971) or branching (e.g. McGuire, 1971) way. He is asked to make diagnostic and management decisions. A variety of technical devices are available to obscure the answers (paper tabs, invisible ink, compounds that can be erased). The pathway of the student through the problem is compared to that of an expert or criterion group and scores are determined for such things as thoroughness, efficiency and proficiency.

The PMP has several virtues. It has content validity and moderate fidelity. It measures, in some form, problem solving, particularly in relation to content oriented data and evaluates skills not tested by objective items. (McGuire & Babbott, 1967). The PMP is now fairly easy to develop, though typically still requires 20–40 hours of work.

The PMP also has limitations. The reliability is unknown and some evidence is available that suggests it may not be appropriate to generalise about problem solving ability from a small number of PMP tests (Page & McGuire, 1973). Although they appear to approximate the clinical process, the medical inquiry study of Elstein et al. (1972) leads one to be wary of using the criterion group method of determining a score. For instance, the more experienced the physician, the lower the thoroughness score tends to be. The only study on criterion related or predictive validity is disturbing. It showed that, in the example chosen, physicians did better on the PMP than they performed in practice and those scoring higher on the PMP did not consistently do best in practice (Goran et al., 1973).

One critical disadvantage is the effect of cueing. For ease of scoring, all PMP's have a pre-determined selection of responses to choose from and research studies have shown that cueing leads to an improved performance (McCarthy, 1966). A modified approach, which avoids cueing has been developed by Berner et al. (1974) – the **Sequential Management Problem (SMP)**. This test does not attempt to evaluate data gathering skills and provides the student with a written history and examination. The student is required to generate problem lists, choose tests and select treatment, indicating the reasons for his decisions. After each section is completed, the student is given feedback before proceeding. A criterion group was used to determine acceptable responses and to provide the feedback information. Another approach is the so-called **Modified Essay Question** recently introduced by Hodgkin & Knox (1975). The validity of these techniques has not been determined but they clearly show great promise.

In conclusion, the PMP or its variants must still be regarded as experimental until more data on reliability and validity is available. It has achieved enthusiastic acceptance by students particularly when used as a teaching tool, with immediate feedback and discussion. It can be expensive and construction requires a lot of time. Scoring is a major problem. Despite the limitations, the technique has great potential for evaluating at least some aspects of clinical performance and is a useful and practical adjunct to objective tests.

5. Computer Simulations

The interest in computer assisted instruction and evaluation (CAI) stems from work in programmed instruction (Skinner, B.F., 1968) and the advances in electronic data processing. An increasing use of CAI is evident in medical education (Stolurow et al., 1970; Hoffer et al., 1975). Technical considerations limits the applicability of CAI and will be briefly mentioned. Any computer programme requires a clearcut statement of objectives which must then be reduced to an algorithm suitable for programming. A particular computer system will only accept information presented in a particular code or language and over 30 of these languages are in use. A language suitable for a simple step by step programme is not usually suitable for developing a programme where a free response between student and computer is envisaged. The computer system will pose limitations, not only in terms of language, but also in terms of the devices which are used to communicate with the student, e.g. typewriter, cathode ray tube, plasma tube, pressure sensitive tube. The more sophisticated systems (e.g. PLATO, IBM1500) allow additional hardware to be incorporated at the terminal and to be computer controlled.

Examples include slide and film projectors, microfiche projectors, random access audio and videotape systems and even vocal response devices (Atkinson and Wilson, 1969; Stolurow, 1970).

Despite the obvious potential of computers and the fact that many medical schools in North America have experimental programmes few have enjoyed widespread application. Apart from the problem of user acceptance (usually Faculty), the issue is predominantly one of cost-efficiency. Time sharing systems are now making wider use more feasible. An alternative to using large fixed computers is to purchase or hire free standing mini-computers.

In undergraduate medical education, the major development of CAI has occurred at three institutions; Ohio State University, Massachusetts General Hospital and the University of Illinois. Ohio State is well known for its Pilot Medical School project which was essentially an independent study programme alternative to the traditional course. The computer played a major role in assisting instruction, evaluating progress and managing the students study (Weinberg, 1973). Ongoing collaboration exists between these institutions and the Lister Hill National Centre for Biomedical Communications at the National Library of Medicine and between them they currently account for over 70 per cent of available educational computer programmes in North America (Brigham & Knapp, 1974).

In regard to the specific area of clinical competence, the computer has provided an alternative to paper and pencil patient management problems. Many examples are available but the most sophisticated is CASE – a Computer Aided Simulation of the Clinical Encounter – initially developed by Harless et al. (1971) at the University of Illinois using the IBM 360–370 computer. This is essentially a branching PMP but with limited natural language capabilities. A major advantage of this is avoidance of the cueing effect of simple PMP's. An associated programme was developed – GENESYS – in which the computer conducts a dialogue with the medical author and semi-automatically generates new cases (Harless et al. 1973). Similar programmes have been developed in other centres (deDombal et al. 1972; Hoffer et al. 1972). A variety of ingenious modifications have been introduced including time factors, test availability and cost (Friedman, 1973).

The most recent developments surround the relatively new computer system, known as PLATO, which was specifically designed for educational use (Bitzer, 1970; Alpert & Bitzer, 1970; Smith & Sherwood, 1976). It requires specific terminal hardware, is operational on the IBM 360 and SDS 940 computers and has its own CAI language (TUTOR). Because of its capacity to incorporate a wide variety of audio-visual displays in its output to the student, it has great flexibility and potential as a teaching and evaluation tool and as a simulation device. The current hope is that this system will become standardised and adopted on a wide basis by many institutions.

The potential of new communications devices, such as satellites, makes it feasible for timeshare networks to be extended overseas. Such networks would potentially increase usage greatly and consequently cut down on cost, not only in terms of user cost but also in terms of development time and cost. It is currently

estimated that a new programme requires 100–200 hours of development time per hour of programme time.

There seems no doubt as to the potential of the computer in the area of evaluating clinical competence. The fidelity of simulation should ultimately be quite high. The capacity to individualise evaluation, perhaps based on a practice profile, would be of particular value in the area of re-certification of practising physicians. Computerised PMP's have already been used by the Royal Colleges of Medicine and Surgery of Canada in a multicentre exam (Osbaldeston, 1974). The present cost of development and operation limits small scale use of the more sophisticated programmes but there is still a great potential for the use of the simpler instructional and evaluation modes as an adjunct to other teaching and testing devices. The use of low cost typewriter and TV display terminals with timesharing facilities makes such use economically feasible.

6. Other Simulations

Simulation may be defined as an instructional method that seeks to provide learning in a truthful and realistic representation of the real world situation in which subsequent independent performance will occur (Maatch, 1972).

Hilliard Jason, one of the doyens of medical education, has stated that he feels simulation is likely to become the most important new educational approach of the decade (Jason and Tichtov, 1971). An alternate viewpoint is stated by Pappworth (1974):

> "God forbid that British teachers and examiners should ape the Americans by employing professional actors to play the part of patients and simulate special disease, or showing still or cine-films instead of actual patients. May we be saved from a new generation of students prodding plastic swellings in dummies and analysing artificial urine. Our profession must be protected from the technomaniacs who consider clinical medicine to be a game which can be played with artificial materials instead of actual patients, and that by manipulating a computer, even a sick one, clinical ability can either be learned or assessed."

This must be regarded as an example of extreme British conservation, anti American feeling, gross ignorance or a total misunderstanding of the aims of those responsible educators who are developing aspects of simulation. While some may be overenthusiastic about particular simulation techniques or devices and certainly not all will be found to be practically useful, the concept of simulation has been proven to be sound.

Simulation has been used extensively by the military since the early 19th Century in war games and military exercises and perhaps achieved its ultimate 20th Century sophistication with the NASA simulations of space flight and the moon environment for astronaut training. Flight simulation for pilot training and evaluation has become standard practice in the airline industry. The 1950's saw a

major development of business and industry games and in the early 1960's the same techniques were applied to education (Tansey and Unwin, 1969; Twelker, 1969; Zuckerman and Horn, 1970). The potential of simulations is being pursued in many medical education centres (Tinning, 1975). Patient management problems, either in paper or computerised format, are in a real sense simulations. However, in this section the emphasis will be on three other areas where there is particular emphasis at present; simulated patients, simulation devices of various types and patient games.

Simulated Patients

Simulated patients have been predominantly used for teaching and evaluating skills in history taking and physical examination and the contribution of Barrows to this field should be acknowledged (Barrows and Abrahamson, 1964; Barrows 1971). He has defined the simulated patient as:

> "a person who has been trained to completely simulate a patient or any aspect of a patient's illness, depending upon educational need. A simulated patient can reproduce faithfully the physiological, emotional, historical and physical manifestations of a patient on observation, interview and examination".

Anyone who has interviewed and examined his simulated patients can be in no doubt as to their ability to match up to this definition. While Barrows has been particularly interested in developing simulations in neurology and internal medicine, albeit with added psychosocial problems, others have found them equally applicable in simulating psychiatric conditions and teaching basic interviewing skills (Helfer and Ealy, 1972; Jason et al., 1971; Adler and Enelow, 1966).

It is claimed, with justification, that simulated patients can often be more effective instructional and learning tools than real patients. They are more readily available and can be scheduled, they can overcome certain risk problems and they can relax the student if he knows he cannot harm the patient. The teaching possibilities can be widened as situations may be presented which would not normally be available or which might otherwise be ethically and medically unacceptable.

The fidelity of the simulation is important, mainly in terms of the psychological fidelity to the student. As long as the simulation appears real it is satisfactory and often this allows the introduction of additional learning tools (e.g. audiovisual aids) which can supplement the simulation without undue loss of fidelity from the learning point of view. This may explain why lower fidelity simulations like role playing, which may be undertaken by the instructor (or examiner) or even a fellow student, may often be acceptable as coming close to a 'real' experience.

A further advantage can be obtained by training the simulator to provide feedback to the student. It is possible to interrupt an interview to discuss any difficulties that arise and then proceed or backtrack and try an alternative

approach. Perhaps the most novel approach to feedback has been the use of prostitutes and, more recently, graduate assistants to teach the routine pelvic examination (Godkins et al., 1974).

With the growing interest in competence based education, it seems likely that more use will be made of simulated patients. This may be particularly so in certification examinations where patient variability has been a major factor in reducing the reliability of clinical examinations. In addition to providing a stable patient problem for testing, a simulated patient can also be trained to make evaluative judgments on the candidates performance including his interpersonal and affective skills. Simulated patients have already been successfully used in the certification examination conducted by the Canadian College of Family Practice (Lamont and Hennen, 1972) to evaluate attitudinal skills. Role playing, a situation where an examiner takes the part of a patient, has been used by the same College and also by the American Board of Orthopaedic Surgery (Levine and McGuire, 1970) and, in Australia, by the Royal Australian College of General Practitioners.

b. Simulation Devices

There are currently a multitude of devices which enable the student to react with models, equipment, machines and audio-visual materials. The majority are directed towards skill training, particularly physical examination skills. Several of the leading medical education centres (e.g. Universities of Southern California and Illinois, Michigan State University) are conducting major research and development programmes in simulation devices and have clinical simulation laboratories for student training (Sajid et al., 1975).

The most famous simulation device is Sim I, a life like computer controlled manikin for testing skills in resuscitation and anaesthesia (Denson and Abrahamson, 1969). Despite a current purchase price of around $100,000, Hoffman and Abrahamson (1975) have established cost-effectiveness and, in addition, have emphasised that when potentially dangerous skills are being taught, a realistic simulation device may have its greatest benefit in reducing patient risk. Even with less dangerous skills, the establishment of a minimum level of skill before patient contact would save many patients considerable discomfort. Simulation devices, even of low fidelity, are well accepted by students in this preparatory phase. Early reports on the effectiveness of simulation devices are encouraging (Penta and Kofman, 1973; Ravin, 1969). Many are now commercially available and the following is an incomplete list of these with developers or manufacturers and approximate costs, where available, in U.S. currency (1975).

Sim I and Sim II: Computer controlled mannequin – resuscitation and anaesthesia, respirator therapy, intravenous therapy, intramuscular injections. Denson and Abrahamson, 1969. ($100,000)

Cardiology Patient Simulator: Animated mannequin with all major cardiovascular signs and capable of simulating more than 50 cardiac disease states. Gordon, 1974. (Under development)

Adult Intubation Simulator: Laerdal Medical Corporation. ($435)

Infant Intubation Simulator: NASCO. ($400)

Bartner Eye Model: Bartner. ($200)

Colenbrander Ophthalmoscopy Mannequin: Ophthalmoscopic Development Laboratory. ($300)

Glaucoma Test Head Model: Richard Rush Studio Inc. ($265)

Keeler Eye Model: Keeler Instruments. ($20)

Strabismus Cover Test Demonstrator: Narco Bio Systems Inc. ($500)

Heart Sound Simulator: Narco Bio Systems Inc. ($9,000)

Cambridge Heart Sound Recorder: Cambridge Instruments Co. ($1,500)

Breath Sound Simulator: Bioinstrumentation Laboratory, University of Illinois (under development).

Ear Model: University of Illinois (under development).

Laryngeal Model: University of Illinois (under development).

Lymph Node/Breast Simulator: University of Illinois (under development).

Breast Simulator (Betsy): Ortho Pharmaceutical Corp.

Ortho Female Pelvis (Gynny): Ortho Pharmaceutical Corp. ($220)

Prostate Palpation Simulator: NASCO. ($400)

Proctoscopy/Sigmoidoscopy Model: Michigan State University (under development).

Spinal Tap Simulator: NASCO (under development).

Resucci Annie: (In general use.)

Infusion Training Arm: Laerdal Medical Corp. ($115)

Injectable Training Arm: NASCO. ($250)

In addition to these specific models and machines, a variety of audio-visual materials are being used as simulation devices. Films and slides have been incorporated into examinations to test observational and interpretive skills (Hubbard et al., 1965; Harden et al., 1975). Films or videotapes of patient

interviews and patient situations have also been extensively used to develop alternatives to direct observation or simulated patients in evaluating interviewing skills (Langsley and Aycrigg, 1970; Cline, 1961). One of the most innovative and successful programmes in teaching interpersonal skills, developed by Kagan and his colleagues (Kagan and Schauble, 1969; Jason, Kagan et al., 1971), uses film and videotape extensively. As with simulated patients, the object of using these methods is to standardise the encounter so that the reliability and validity of the testing procedure can be determined.

c. Patient Games

A game can be looked on as a model which is an abstraction of things held to be important in real life (Tansey and Unwin, 1969). It will have certain characteristics depending on how well it represents the real life situation (fidelity), how much of the real life situation is covered and how easy it is to understand the model and the principles it incorporates. Games can be as simple or as complex as desired and have been used extensively in schools (usually simple) and by the military and business (usually complex) where the ability to simulate complex and potentially dangerous situations (e.g. wars, take-overs) has important advantages. Games have also been developed which teach social and moral concepts (Tansey and Unwin, 1969; Zuckerman and Horn, 1973).

Though PMP's could be regarded as patient games, they lack the personal involvement that the games described above imply. Game playing is usually undertaken in a group situation and competition may or may not be introduced as a factor. Motivation is invariably high. Despite an obvious potential for use in medical education, little work has been done in this field. Maatch (1974) has developed a Patient Game which can be played by individuals or groups. The game is designed to teach and evaluate the fundamentals of diagnosis and management. The instructor is provided with a kit containing information about the patient, which the instructor plays, and a flow diagram on which he charts the students decisions. He provides feedback on examination, laboratory data, consultations, treatment and procedures, including time, costs and effect on the patient. A high degree of fidelity can be achieved even when simulating emergency situations. Evaluation is undertaken by comparing actual performance against key instructional objectives. A more sophisticated computerised game has been developed, under the auspices of the British Pharmaceutical Industry, in which the participants conduct a clinical trial on a new drug. Over a three-day workshop period, a research design is prepared and the experiment conducted and evaluated (Maxwell, Smith and Shing, 1972).

A simpler game is being developed by Barrows at McMaster University which can be used as a self-learning device or used in tutorial groups. It consists of a deck of cards starting with a brief history and allows the student to work through the history, examination, investigations and consultations in any way or order he wishes. The student is constantly challenged to justify his decisions. The card deck is included in a "problem box" with a variety of additional material which may be required (Barrows & Mitchell, 1975). The McMaster group are also working on a

medical take-off of Monopoly and have computer programmes which allow complex physiological experiments and games to be undertaken on "normal" and "abnormal" patients (Dickinson, 1973). These programmes are known as MacMan (cardiovascular), MacPuff (respiratory) and MacPee (fluid balance-electrolytes).

Summary

There seems little doubt that simulations, appropriately designed, can induce a high degree of motivation which is one of the prime prerequisites for learning. This may be particularly appropriate when it is difficult to provide "real" situations to deal with. There is no direct evidence that the method itself improves the transfer of learning (Cherryholmes, 1966) and clearly the ultimate goal is to establish whether simulated experiences transfer to similar tasks in the real world (Gagne, 1962). However, where used in the fields of intercultural communication, counselling and administrative decision making and problem solving, studies have shown a significant change in attitudes and improvement in performance. This is particularly encouraging for medicine where these issues are of current concern.

7. Oral Examinations

The oral or clinical viva examination has for centuries been the pre-eminent method, and sometimes the only method, used for evaluating a student's clinical performance at the final certifying examination. In British and Australian medical schools, it still retains a hallowed place and is similarly prominent in specialist college examinations. There is no doubt that the oral examination has features which are not present in any other examination format, notably the examiner – examinee interaction. Unfortunately, the traditional oral examination, in medical and other fields of education, has been shown to have severe deficiencies in terms of reliability (Ebel, 1972; Mehrens and Lehmann, 1973; Hubbard, 1971). The use of oral examinations for certification purposes is under widespread criticism, particularly in North America, as the reliability of a variety of well conducted specialist board examinations has been shown to be unacceptable (Foster et al., 1969; McGuire, 1966; Pokorny and Frazier, 1966). A large study of the bedside oral examination conducted by the National Boards of Medical Examiners showed that the correlation between examiners, when evaluating the same student on different cases, was no greater than chance (Hubbard et al., 1965).

McGuire (1966) in her study of over 2,000 oral examinations showed that the examiners were testing, predominantly, the recall of isolated fragments of information and Evans et al., (1966) demonstrated that 76% of questions in a medical student examination in internal medicine were at the recall level. One of the major problems is to reduce inter-examiner variability but efforts to improve this with training have not been particularly successful (Ludbrook and Marshall, 1971).

These studies have led to two reactions. The first followed, for example, by the National Boards (Hubbard et al., 1971) and the American Board of Internal

Medicine (ABIM, 1971) has been to discontinue oral examinations. The second, and more constructive, has been to recognise that oral examinations can be used to measure aspects of clinical competence not easily measured by other techniques providing these aspects are clearly defined (McGuire, 1966). This approach is reflected in the examination conducted by the American Board of Orthopaedic Surgery (Levine and McGuire, 1970a) and more recently by the Canadian College of Family Physicians (Van Wart, 1974). Measures to improve reliability have included the use of role playing or simulated patients to reduced patient variability, structured questioning and examiner training programmes (Levine and McGuire, 1970b; Van Wart, 1974). Similar efforts need to be made at the medical student level.

8. Direct Observation

In recent years interest has developed in evaluating performance skills in ward and practice settings and consequently it has become necessary to observe the student as he performs clinical tasks (Hinz, 1966). It is remarkable how infrequently in the past the student has been systematically observed during these student–patient interactions. Direct observation techniques include check lists, rating scales, participation charts and anecdotal records. These are deceptively easy to develop and may have high face or content validity, particularly if the criterion to be evaluated have been clearly defined. Strictly speaking, check lists, participation charts and anecdotal records should not include subjective interpretation at the time of observation and therefore are potentially more reliable.

Rating scales imply judgment in a systematic manner and unless carefully constructed and scored are open to many errors. These errors include the ambiguity of the scale itself, the halo effect, leniency or severity effects and errors of central tendency. The validity and reliability of the rating can be increased by improving the rating scale, improving the rater or improving the rating method. This is a subject with which major textbooks on measurement in education and psychology deal at length (Thorndike and Hagen, 1969; Guildford, 1954; Ebel, 1972; Mehrens and Lehmann, 1973). A comprehensive review of the area, paying particular attention to new techniques of potential value in the evaluation of clinical performance, has been prepared by Fink et al., (1974) for the National Board of Medical Examiners (over 400 references). The Board is currently looking at the development of rating forms for clinical evaluation of physician's assistants (Dowalby and Andrew, 1976).

The distinction has been drawn between technical skills and interpersonal skills and it is recognised that both are important and interact in the clinical situation (Jason, 1970). The emphasis in both undergraduate and postgraduate education has been on evaluating technical skills and in recent years some medical schools have begun to develop programmes to teach and evaluate interpersonal and communication skills. Notable examples of this are being conducted at Michigan State University and McMaster University where several new methods of assessment are in use (Werner and Scheider, 1974; Finkel and Norman, 1973; Norman et al., 1975). Other approaches are described in the review by Barro (1973).

Most direct ob ervation assessments are made by teachers but Helfer (1972) discusses the role of student peer evaluation and points out the potential of peer ratings for measuring aspects of interpersonal skill and attitudes which are not readily observed by a teacher. Although it seems unlikely that direct observation will be superceded as the prime method for assessing interpersonal skills in the clinical setting, there have been some interesting developments using multiple choice questions to evaluate reactions to filmed interviews or to a series of stated doctor–patient situations (Werner and Schneider, 1974; Rasche et al., 1973; Langsley and Aycrigg, 1970).

At present, it is not possible to make generalisations about the use of direct observation techniques in evaluating clinical competence. Because of the difficulties in obtaining reliable scores, they should probably be mainly used for summative (in-course) evaluation. This need not apply if the examination is carefully structured, the aims clearly defined and the examiners are prepared to undergo training to assure satisfactory reliability. The use of simulated patients, in some settings, should prove to be an added advantage. If important decisions are to be made on ratings, then expert psychometric advice should be sought.

9. Record Audit

Medical records are rarely standardised, frequently incomplete and often inaccurate (Gonnella et al., 1970; Donabedian, 1967) and there is no clearcut evidence that good records equate with good patient care (Helfer, 1967; Scott and Sniderman, 1973; Rosenfeld, 1957). Students' case write-ups or ward charts are commonly evaluated and are often superior to the regular hospital charts but there is no reason to assume that the chart quality reflects the quality of patient care unless combined with direct observation. The problem oriented medical record developed by Weed (1968), and advocated since by many others, is, potentially, a useful way of standardising the record so that more reliable judgments on performance can be made (Hurst, 1971; Newble, 1973). Difficulties in implementing the system have often been encountered (Newble et al., 1974).

Record auditing is being widely investigated as a potential measure of performance in medical practice both in hospital and in private practice. The emphasis on this comes from the North American concern for developing measures of quality control of patient care and about recertification. The review by Barro (1973) is a readily available resource document in this area.

10. Multiple Methods

It should be clear from the foregoing discussion that no one testing method is capable of giving a reliable, valid or practical evaluation of all aspects of clinical competence. The critical incident studies conducted by the NBME to define clinical competence led to the concept of using several test methods during one examination and trying to match the most appropriate method to the aspect of clinical performance being tested (Hubbard et al., 1965). Some of the new methods

introduced into the Board's examinations have since been abandoned because of the practical difficulties in implementing them in such large numbers or because they failed to reach the required level of objectivity. The same reasoning led to the cessation of the bedside oral examination.

A different approach has been adopted by the American Board of Orthopaedic Surgery who have developed a multiple method format and the same methodology has been introduced by the Canadian College of Family Practice. A similar but more ambitious testing scheme is projected for the American College of Emergency Physicians. These examinations, which are discussed in more detail in the next chapter, include, for example, the use of multiple choice questions, paper patient management problems, simulated patients, role playing oral examinations and direct assessment. A profile system is used for scoring (Levine and McGuire, 1971).

These examinations reflect a major input by two of the leading medical education institutions in America. Medical student teaching has not reflected this global approach to evaluating clinical performance probably because of the pre-eminent place of the NBME in North America in the certification of clinical competence. The position of British and Australian medical schools is more flexible. An objective structured examination conducted at Dundee, in which students rotate round a series of stations and are evaluated on several aspects of clinical competence, is a valuable model for further development (Harden et al., 1975). Moves toward multiformat examinations are also evident in Australia (Isbister and Egerton, 1976).

The multi method examinations discussed above seem to be the growing trend and most likely to be implemented on a wide scale in the immediate future. They are all "process" approaches. An entirely different "qualities approach" has come out of the work of Price et al. (1971) at the University of Utah. They have identified, using the factor analysis technique, 80 criteria of the 'good' and 'bad' physician. They have used a variety of resources to identify these criteria including doctors, students, nurses, hospital administrators, the general public and patients. The criteria have been ranked and rating forms developed which can be used to evaluate the performance of individual physicians. The ratings have been shown to be reliable but the validity has not been established. Despite this, several recent reviewers of the field have identified this method as having great potential. There is reason to believe from work in other fields that a qualities approach is more likely to be successful in predicting future performance (Barro, 1973).

11. Outcome Measures

It has been mentioned repeatedly that process methods are only valid to the extent to which they predict outcome. The outcome measures vary in emphasis. They may be concerned with final or intermediate consequences, with long term or short term results, with generality or with particular diseases and problems. Reliable outcome measures are rarely available and do not seem to have been used in evaluating undergraduate students. This is hardly surprising as few students are in a

position to influence outcome. However, some of the intermediate outcome measures could be developed for student assessment. The main thrust in this area is towards measuring the clinical performance of practising physicians and the reader is referred for further information to the review by Barro (1973) and to the work of Williamson (1971), Sanazaro and Williamson (1968) and Brook and Appel (1973).

REFERENCES

Introduction to Methods

Barro, A.R., Survey and Evaluation of Approaches to Physician Performance Measurement. J. Med. Educ. 48 1050, 1973.

Bloom, B.S., (Ed) Taxonomy of Educational Objectives: Cognitive Domain. David McKay Co. Inc. 1956.

Dudley, H.A.F., Taxonomy of Clinical Educational Objectives. Brit. J. Med. Educ. 4 13, 1970.

Engel, G.L., Care and Feeding of the Medical Student. The Foundation for Professional Competence. J.A.M.A. 215 1135, 1971.

Engel, G.L., Enduring Attributes of Medicive Relevant for the Education of the Physician. Ann. Int. Med. 78 587, 1973.

Feinstein, A.R., What Kind of Basic Science for Clinical Medicine. N.E.J.M. 283 847, 1970.

Gagne, R.M., The Conditions of Learning (2nd edition).

Holt, Rinehart and Winston Inc. 1970.

Harrow, A.J., A Taxonomy of the Psychomotor Domain. David McKay Co. Inc. 1972.

Irby, D.M. and Morgan, M.K., Clinical Evaluation: Alternatives for Health Related Educators. The Centre for Allied Health Instructional Personnel. Gainesville, Florida, 1974.

Kane, R., Ross Woolley, F., and Kane, R., Towards Defining the End Product of Medical Education. J. Med. Educ. 48 615, 1973.

Krathwohl, D.R., Bloom, B.J. and Masia, B.B., Taxonomy of Educational Objectives: Affective Domain. David McKay Co. Inc. 1964.

Sanazaro, P.J. and Williamson, J.W., Physician Performance and its Effect on Patients. Medical Care. 8 299, 1970.

Stead, E.A., Physicians – Past and Future. Arch. Int. Med. 127 703, 1971.

Williamson, J.W. Evaluating Quality of Patient Care: A Strategy Relating Outcome and Process Assessment. J.A.M.A. 218 564, 1971.

Wingard, J.R. and Williamson, J.W. Grades as Predictors of Physicians Career Performance: An Intuative Literature Review. J. Med. Educ. 48 311, 1973.

Test Characteristics

Adkins, D. Test Construction (2nd edition). Charles Merrill 1974.

Hubbard, J.P. Measuring Medical Education. Lea and Febiger 1971.

Essay

Adkins, D., Test Construction (2nd edition). Charles Merrill 1974.

Completion or Short Answer

Adkins, D., Test Construction (2nd edition). Charles Merrill 1974.

Newble, D. and Baxter, A., Unpublished observations 1976.

Objective (Multiple Choice)

Adkins, D., Test Construction (2nd edition). Charles Merrill 1974.

Bloom, B.S. (Ed.), Taxonomy of Educational Objectives: Cognitive Domain. David McKay Co. Inc. 1956.

Hubbard, J.P., Measuring Medical Education. Lea and Febiger 1971.

Levine, H.G., McGuire, C.H. and Nattress, L.W., The Validity of Multiple Choice Achievement Tests as Measures of Competence in Medicine. Amer. Educ. Research J. 7 69, 1970.

Wingard, J.R. and Williamson, J.W., Grades as Predictors of Physicians Career Performance: An Evaluative Literature Review. J. Med. Educ. 48 311, 1973.

Paper Simulations

Berner, E.S., Hamilton, L.A. and Best, W.R., A New Approach to Evaluating Problem Solving in Medical Students. J. Med. Educ. 49 666, 1974.

Elstein, A., Kagan, N., Shulman, L., Jason, H. and Loupe, M.J., Method and Theory in the Study of Medical Inquiry. J. Med. Educ. 47 85, 1972.

Galofre, A., A Review of Written Paper Management Simulations (unpublished), Center for Educational Development, University of Illinois, 1974.

Goran, M.J., Williamson, J.W. and Gonnella, J.S., The Validity of Patient Management Problems. J. Med. Educ. 48 171, 1973.

Helfer, R.E. and Slater, C.H., Measuring the Process of Solving Clinical Diagnostic

Problems. Brit. J. Med. Educ. 5 48, 1971.

Hodgkin, K. and Knox, J.D.E., Problem Centred Learning: The Modified Essay Question in Medical Education. Churchill-Livingstone 1975.

Hubbard, J.P., Measuring Medical Education. Lea and Febiger, 1971.

McCarthy, W.H., An Assessment of the Effect of Cueing Items in Objective Examinations. J. Med. Educ. 41 263, 1966.

McCarthy, W.H. and Gonnella, J.S., The Simulated Patient Management Problem: A Technique for Evaluating and Teaching Clinical Competence. Brit. J. Med. Educ. 1 348, 1967.

McGuire, C.H. and Babbott, D., Simulation Technique in the Measurement of Problem Solving Skills. J. Med. Educ. 4 1, 1967.

McGuire, C. and Solomon, L.M., Clinical Simulations. Appleton, Century and Crofts. 1971.

Page, G. and McGuire, C.H., An Overview of Physicians Performance on Simulated Clinical Encounters Presented in PMP Format. Proceedings of AERA. 1973.

Rimoldi, H.J.A., Rationale and Application of the Test of Diagnostic Skills. J. Med. Educ. 38 364, 1963.

Computer Simulations

Alpert, D. and Bitzer, D.L., Advances in Computer Based Education. Science 167 1582, 1970.

Atkinson, R.C. and Wilson, H.A., Computer Assisted Instruction. Academic Press 1969.

Bitzer, D.L., PLATO in Computer Assisted Instruction in The Health Professions. Entelek Inc. 1970.

Brigham, C.R. and Kamp, M., Current Status of Computer Assisted Instruction in Health Sciences. J. Med. Educ. 49 278, 1974.

de Dombal, F.T., Smith, R.B., Modgill, V.K. and Leaper, D.J., Simulation of the Diagnostic Process: A Further Comparison. Brit. J. Med. Educ. 6 238, 1972.

Friedman, R.B., A Computer Program for Simulating the Patient Physician Encounter. J. Med. Educ. 48 92, 1973.

Harless, W.D., Drennon, G.G., Marxer, J.J., Root, J.A. and Miller, G.E., CASE: A Computer Aided Simulation of the Clinical Encounter. J. Med. Educ. 46 443, 1971.

Harless, W.D. et al., GENESIS: A Generating System for the CASE Natural Language Model. Comput. Biol. Med. 3 Pergammon Press 1973.

Hoffer, E.P., Barnett, G.O., Farquhar, B.B. and Prather, P.A., Computer-Aided Instruction in Medicine. Annu. Rev. Biophys. Bioeng. 4 103 1975.

Hoffer, E.P., Barnett, G.D. and Farquhar, B.B., Computer Simulation Model for Teaching Cardio-Pulmonary Resuscitation. J. Med. Educ. 48 92, 1973.

Osbaldeston, L.W., CBI in Medicine at the University of Alberta. J. of Comp. Based Instr. 1 64, 1974.

Skinner, B.F., The Technology of Teaching. Appleton-Century-Crofts, 1968.

Smith, S.G. and Sherwood, B.A., Educational issues of the PLATO computer system. Science, 192, 344, 1976.

Stolurow, L.M., Peterson, T.I. and Cunningham, A.C., Computer Assisted Instruction in the Health Professions. Entelek Inc. 1970.

Weinberg, A.D., CAI at the Ohio State University College of Medicine. Comput. Biol. Med. 3 299, Pergammon Press 1973.

Other Simulations

Adler, L.M. and Enelow, A.J., An Instrument to Measure Skill in Diagnostic Interviewing. A Teaching and Evaluation Tool. J. Med. Educ. 41 281, 1966.

Barrows, H.S. and Abrahamson, S., The Programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology. J. Med. Educ. 39 802, 1964.

Barrows, H.S., Simulated Patients (Programmed Patients). Charles C. Thomas 1971.

Barrows, H.S. and Mitchell, D.L., An Innovative Course in Undergraduate Neuroscience. Experiment in Problem-Based Learning with 'Problem Boxes'. Brit. J. Med. Educ. 9 223, 1975.

Cherryholmes, C.H., Some Current Research on Effectiveness of Educational Simulations. Amer. Behav. Scientist, 10 4, 1966.

Cline, M.G., A Film Test of Clinical Skills in Medical Students. J. Med. Educ. 36 908, 1961.

Denson, J.S. and Abrahamson, S., Computer Controlled Patient Simulation. J.A.M.A. 208 504, 1969.

Dickinson, C.J., Digital Computer Model of Gas Exchange and Blood Circulation for Research and Teaching. Proc. R. Soc. Med. 66 392, 1973.

Gagne, R., Simulators. In Glaser (Ed) Training Research and Education. University of Pittsburg Press, 1962.

Godkins, T.R., Duffy, D., Greenwood, J. and Stanhope, W.D., Use of Simulated Patients to Teach the 'Routine' Pelvic Examination. J. Med. Educ. 49 1174, 1974.

Gordon, M.S., Cardiology Patient Simulator. Amer. J. Cardiol. 34 350, 1974.

Harden, R. McG., Stevenson, M., Downie, W.W. and Wilson, G.M., Assessment of Clinical Competence using an Objective Structured Examination. Brit. Med. J. I 447, 1975.

Helfer, R.E. and Ealy, K.F., Observations of Pediatric Interviewing Skills. Amer. J. Dis. Child. 123 556, 1972.

Hoffman, K. and Abrahamson, S., Personal Communication (to be published), 1975.

Hubbard, J.P., Levit, E.J., Schumacher, C.F. and Schnabel, T.G., An Objective Evaluation of Clinical Competence. N.E.J.M. 272 1321, 1965.

Jason, H. and Tichtov, S. (Eds.), Instructional Technology in Medical Education. Vol. II, R.R. Bowker and Co., 1971.

Jason, H., Kagan, N., Werner, A., Elstein, A. and Thomas, J., New Approaches to Teaching Basic Interview Skills to Medical Students. Amer. J. Psych. 127 1404, 1971.

Kagan, N. and Schauble, P.G., Affective Simulation in Interpersonal Process Recall. J. Counseling Psychol. 16 309, 1969.

Lamont, C.T. and Hennen, B.K.E., The Use of Simulated Patients in a Certification Examination in Family Medicine. J. Med. Educ. 47 789, 1972.

Langsley, D.G. and Aycribb, J.B., Filmed Interviews for Testing Clinical Skills. J. Med. Educ. 45 52, 1970.

Levine, H.G. and McGuire, C.H., The Use of Role Playing to Evaluate Affective Skills in Medicine. J. Med. Educ. 45 700, 1970.

Maatch, J.L., Unpublished Grant Proposal. 1972.

Maatch, J.L., An Introduction to Patient Games: Some Fundamentals of Clinical Instruction. Biomed. Comm. Center, Michigan State Univ., 1974.

Maxwell, C., Smith, F.R. and Shing, L.K., Instant Experience in Clinical Trials – A Computerised Teaching Aid. Trust for Education and Research in Therapeutics. 1972.

Pappworth, M.H., Passing Medical Examinations. Butterworths, 1974.

Penta, F.B. and Kofman, S., Effectiveness of Simulation Devices in Teaching Selected Skills of Physical Diagnosis. J. Med. Educ. 48 442, 1973.

Sajid, A., Lipson, L.F. and Telder, T.V., A Simulation Laboratory for Medical Education. J. Med. Educ. 50 970, 1975.

Tansey, P.J. and Unwin, D., Simulation and Gaming in Medical Education. Methuen Educational Ltd., 1969.

Tinning, F.C., Simulation in Medical Education. Michigan State Univ., 1975.

Twelker, P.A., Instructional Simulation: A Research and Dissemination Activity: Oregon State University Continuing Education and ERIC Publication. 1969.

Zucherman, D.W. and Horn, R.E., The Guide to Simulations/Games for Education and Training. Information Resources Inc., 1973.

Oral Examinations

American Board of Internal Medicine. Written (Certifying) Examination in Internal Medicine. Bull Amer. Coll. Phys., 12 236, 1971.

Ebel, R.L., Essentials of Educational Measurement. Prentice-Hall, 1972.

Evans, L.R., Ingersoll, R.W. and Smith, E.J., The Reliability, Validity and Taxonomic Structure of the Oral Examination. J. Med. Educ., 41 651, 1966.

Foster, J.J., Abrahamson, S., Lass, S. and Girard, M.A., An Analysis of an Oral Examination used in Specialty Board Certification. J. Med. Educ., 44 951, 1969.

Hubbard, J.P., Measuring Medical Education. Lea and Febiger, 1971.

Hubbard, J.P., Levit, E.J., Schumacher, C.G. and Schnabel, T.G., An Objective Evaluation of Clinical Competence. N.E.J.M. 272 1321, 1965.

Levine, H.G. and McGuire, C., (a) The Validity and Reliability of Oral Examinations in Assessing Cognitive Skills in Medicine. J. Educ. Meas. 7 63, 1970. (b) The Use of Role Playing to Evaluate Affective Skills in Medicine. J. Med. Educ., 45 700, 1970.

Ludbrook, J. and Marshall, V.R. Examiner Training for Clinical Examinations. Brit. J. Med. Educ. 5 152, 1971.

McGuire, C.H., The Oral Examination as a Measure of Professional Competence. J. Med. Educ. 41 267, 1966.

Mehrens, W.A. and Lehmann, I.J., Measurement and Evaluation in Psychology and

Education. Holt, Rinehart and Winston Inc., 1973.

Pokorny, A.D. and Frazier, S.H., An Evaluation of Oral Examinations. J. Med. Educ. 41 28, 1966.

Van Wart, A.D., A Problem Solving Oral Examination for Family Medicine. J. Med. Educ. 49 673, 1974.

Direct Observation

Barro, A.R., Survey and Evaluation of Approaches to Physician Performance Measurement. J. Med. Educ. 48 1050, 1973.

Dowalby, F.J. and Andrew, B.J., Relationship Between Clinical Competence Ratings and Examination Performance. J. Med. Educ. 51 181, 1976.

Ebel, R.L., Essentials of Educational Measurement. Prentice-Hall, 1972.

Fink, C.D. et al., Review of Techniques and Procedures for the Measurement of Complex Human Performance. Consulting Report by Human Resources Research Organisation for the N.B.M.E., 1974.

Finkel and Norman, G.R., Validity of Direct Observation in Assessment of Clinical Skills. Presented at R.I.M.E. Conference, 1973.

Guildford, J.B., Psychometric Methods (2nd Ed.). McGraw-Hill, 1954.

Helfer, R.E., Peer evaluation: Its Potential Usefulness in Medical Education. Brit. J. Med. Educ. 6 224, 1972.

Hinz, C.F., Direct Observation as a Means of Teaching and Evaluating Clinical Skills. J. Med. Educ. 41 150, 1966.

Jason, H., The Relevance of Medical Education to Medical Practice. J.A.M.A. 212 2092, 1970.

Langsley, D.G. and Aycrigg, J.B., Filmed Interviews for Testing Clinical Skills. J. Med. Educ. 45 52, 1970.

Mehrens, W.A. and Lehmann, I.J., Measurement and Evaluation in Psychology and Education. Holt, Rinehart and Winston Inc., 1973.

Norman, G.G., Wakefield, J.G. and Pineo, G., Experience with a Single Encounter Assessment of Clinical Skills. Submitted for publication, 1975.

Rasche, L.M., Bernstein, L. and Veenhuis, P.E., Evaluation of a Systematic Approach to Teaching Interviewing. J. Med. Educ. 49 589, 1974.

Thorndike, R.L. and Hagan, E., Measurement and Evaluation in Psychology and

Education (3rd Ed.). Wiley, 1969.

Werner, A. and Schneider, J.M., Teaching Medical Students Interactional Skills. N.E.J.M. 290 1232, 1974.

Record Audit

Barro, A.R., Survey and Evaluation of Approaches to Physician Performance Measurement. J. Med. Educ. 48 1050, 1973.

Donabedian, A., A Guide to Medical Care Administration Vol. II Medical Care Appraisal. Amer. Public Health Assoc. Inc. 1969.

Gonnella, J.S., Goran, M.J., Williamson, J.W. and Cotsonas, N.J., Evaluation of Patient Care: An Approach. J.A.M.A. 214 2040, 1970.

Helfer, R.E., Estimating the Quality of Patient Care in a Pediatric Emergency Room. J. Med. Educ. 42 244, 1967.

Hurst, J.W., How to Implement the Weed System. Arch. Int. Med. 128 456, 1971.

Newble, D.I., Teaching Basic Clinical Skills and Problem Solving to Fifth Year Medical Student. Med. J. Aust. 2 1094, 1973.

Newble, D.I., Judd, S.J. and Wangel, A.G., Quality Control of Patient Care – The Practical Application of Problem Oriented Medical Records. Aust. N.Z. J. Med. 4 1050, 1973.

Rosenfeld, L.S., Quality of Medical Care in Hospitals. Amer. J. Public Health. 47 856, 1957.

Scott, H.M. and Sniderman, A., Evaluation of Clinical Competence Through a Study of Patient Records. J. Med. Educ. 48 832, 1973.

Weed, L.L., Medical Records that Guide and Teach. N.E.J.M. 278 593 & 652, 1968.

Multiple Methods

Barro, A., Survey and Evaluation of Approaches to Physician Performance Measurement. J. Med. Educ. 48 1050, 1973.

Harden, R.McG., Stevenson, M., Downie, W.W. and Wilson, G.M., Assessment of Clinical Competence Using Objective Structured Examinations. Brit. Med. J. 1 447, 1975.

Hubbard, J.P., Levit, E.J., Schumacher, C.F. and Schnabel, T.G., An Objective Evaluation of Clinical Competence. N.E.J.M. 272 1321, 1965.

Isbister, W.H. and Egerton, W.S., An Improved M.B.,B.S. Final Examination in Surgery. Med. J. Aust. 1 351, 1976.

Levine, H.G. and McGuire, C.H., Use of a Profile System for Scoring and Reporting Certifying Examinations in Orthopaedic Surgery. J. Med. Educ. 46 78, 1971.

Price, et al., Attributes of a Good Practising Physician. J. Med. Educ. 46 229, 1971.

Price, P., Taylor, C. et al., Measurement of Physician Performance: Two Decades of Intermittently Sustained Research. LLR Press. 1971.

Outcome Measures

Barro, A., Survey and Evaluation of Approaches to Physician Performance Measurement. J. Med. Educ. 48 1050, 1973.

Brook, R.H. and Appel, F.A., Quality of Care Assessment: Choosing a Method for Peer Review. N.E.J.M. 288 1323, 1973.

Sanazaro, P.J. and Williamson, J.W., End Results of Patient Care: A Provisional Classification Based on Reports by Internists. Medical Care. 6 123, 1968.

Williamson, J.W., Evaluating Quality of Patient Care: A Strategy Relating Outcome and Process Assessment. J.A.M.A. 218 564, 1971.

4. EXAMINATIONS OF CLINICAL COMPETENCE

Most members of a University Faculty of Medicine would agree with the contention that the main aim of an undergraduate medical training is to produce a reasonably competent doctor. Unfortunately, few have defined the entity "competent doctor" and what is "reasonable" in terms of performance standards at the point of graduation. This problem is not confined to medicine and has led to major efforts by educationalists to define objectives which give a clear indication of what a student should be able to do at the completion of his training (behavioural and terminal objectives) (Mager, 1967; Thorndike, 1971).

Criterion Referenced Versus Norm Referenced Examination

In the professional field, the recent concern has been for the development of criterion referenced measurements (Glaser and Nitko, 1971; Popham, 1969, 1974). Such measurements entail a clearcut statement of the minimum acceptable competence of an individual over the range of knowledge, skills and attitudes which have been identified as necessary in a graduating student. An essential ingredient is the determination of the minimum acceptable performance prior to the examination. This contrasts sharply with the more traditional procedure of norm referenced measurement. In this case the pass or cut off level is predetermined statistically, e.g. one or two standard deviations below the mean, and the students essentially compete with each other. Items for the examination are often chosen for their discriminatory value as much as for their content and there is always a fixed failure rate. Test specialists continue to argue the philosophical issues (e.g. Ebel, 1971) and certain technical problems in establishing reliability of criterion-referenced tests have yet to be solved. The major difficulty in criterion-referenced testing is to determine the minimum pass score as this requires judgment rather than being a purely statistical decision. Often very arbitrary standards are set. It is hoped to overcome these difficulties by using external criterion groups to set pass levels and by developing new test statistics.

Current State of Development of Certification Examinations

Few comprehensive studies of clinical certification examination procedures are available outside North America. A review of the examinations of the National Board of Medical Examiners (NBME) and various specialty boards reveals a wide range of test formats. The NBME (Part III) continues to use a norm

referenced examination and relies on a combination of objective pencil and paper tests and PMP's. Bedside clinical orals were eliminated because of demonstrated unreliability and unfairness (Beck, 1972). In the NBME a fixed percentage failure rate is determined and this is justified on the grounds that there is little variance in performance standard from year to year. However, the NBME is well aware that the examination format does not test all aspects of clinical performance and has an active research programme to determine feasible alternatives.

The American Board of Orthopaedic Surgery (ABOS) was the first organisation to plan and implement a competency based clinical examination. This was done with the guidance of the Centre for Educational Development at the University of Illinois. As this is the prototype for such examinations, its development will be described in some detail. The first step undertaken was to conduct a critical incident study and over 1,000 orthopaedic surgeons participated in this exercise (Blum and Fitzpatrick, 1968). The areas of competence were classified (Miller, 1968). On reviewing the current examination they found that the written examination tested subject matter in an incomplete and spotty manner and that the oral examination largely duplicated the content and process of the written examination (McGuire, 1966). A decision was made to develop an examination which would shift attention to areas of competence other than factual recall.

A series of task forces was set up. The first developed a blueprint of the certifying process which outlined subject matter content, intellectual processes, professional skills and attitudes and put weightings on their relative importance. A second task force classified multiple choice test items and the third created new items where gaps were identified. The fourth task force developed a series of patient management problems (PMP) to test problem solving skills and the fifth was able to achieve agreement among Board examiners to direct the oral examinations towards problem solving, observation and interpretation and interpersonal communication skills. To improve reliability, standardised case materials were available to all examiners and this material was coupled with rating scales and a period of examiner training in the new techniques. Finally, it was recognised that any formal examination could only sample a small area of performance skills and a sixth task force set about developing rating scales to be used over a prolonged period of time prior to the examination.

No single test gives information that is reliable enough to be considered critical in assessing competence. It was therefore impossible to justify failing a candidate on the results of a single test. The scoring system relating to test type was replaced with a profile of performance elements under the following series of abilities: to recall information; to observe, analyse and interpret data; to solve clinical problems; to communicate effectively with patients and colleagues; to exhibit surgical skill. For example, "the ability to solve clinical problems" is assessed from the oral examination and the PMP's. The final result is a battery of test procedures leading to a profile score (Nattress, 1969, 1971; Levine and McGuire, 1971). The results of the first examination held in 1968 showed that many problems had to be solved with this approach but significant deficiencies were identified in candidates which has led to changes in orthopaedic training programmes.

The Canadian College of Family Practice has modelled its examination on the Orthopaedic Boards. It consists of a 220 item multiple choice test, a film strip examination, PMP's and a role playing structured oral conducted in a simulated doctor's office to assess attitudes. Simulated patients are also used. Technical skills to a standard satisfying the candidate's preceptors were a pre-requisite to sitting the examination (Corley, 1971; Lamont, 1972; Van Wart, 1974).

The most ambitious development in criterion referenced testing is that currently being developed by the American College of Emergency Physicians in association with the Office of Medical Education, Research and Development (OMERAD) at Michigan State University. An extensive proposal to develop and evaluate this procedure as a model for all specialty boards has been submitted. Once again a multiple test format has been chosen with an extensive battery of evaluation techniques being incorporated. They propose a two-day examination preceded by a self-scoring diagnostic pretest. The examination consists of an objective test of factual knowledge, an objective test of perceptual and interpretative skills utilising slides and stimulus films, PMP's to test problem solving, simulated patient encounters to evaluate clinical, manipulative and interpersonal skills and simulated situation encounters to test the multiple patient – multiple skills situations required of emergency physicians (Shulman, Maatch et al., 1975).

Implications for Australian Medical Schools

These studies clearly have a major implication for Australian medical schools which have the direct responsibility to certify the competence of students and cannot rely on the services of a national examination centre. The faculty of the school should define and publish the terminal objectives pertinent to clinical competence so that they are available to staff and students. The identification of such objectives will enable a more logical and relevant approach to the setting of final (certification) examinations. This would inevitably have a salutory effect on teaching strategies and student learning. It is obvious to many that the objectives and format of the examination exert a strong motivational force on students and if the objectives are inappropriate then this will lead to a distortion of their learning. At the University of Adelaide, the replacement of the final clinical viva for all but a few students, and the substitution of continuous (progressive) assessment, may have altered student attitudes to the development of practical clinical skills. The indication of graduation success is currently evaluated on the performance in a norm referenced multiple choice examination predominantly requiring recall of factual information. The substitution of a multiple format examination, based on clearly defined objectives and aiming for acceptable reliability and validity, could significantly improve this situation. The next chapter will be concerned with the development of a model that could be utilised for developing such an examination.

REFERENCES

Beck, J.C., Certification and the Specialty Board. Bull. N.Y. Acad. Med. 50 11, 1222, 1972.

Blum, J.M. and Fitzpatrick, R., Critical Performance Requirements for Orthopaedic Surgery. Amer. Inst. Research, 1968.

Corley, J.B., The Evaluation of Clinical Competence in the Family Physician. In Evaluation in Medical Education, Gilbert (Ed) Bulletin Commercial, Edmonton, 1971.

Ebel, R.L., Criterion Referenced Measurements. Limitations. School Review, 79 282, 1971.

Glaser, R. and Nitko, A.J., In Educational Measurement. Thorndike (Ed), American Council on Education, 1971.

Lamont, C.T. and Hennen, B.K.E., The Use of Simulated Patients in a Certification Examination in Family Medicine. J. Med. Educ. 46 789, 1972.

Levine, H.G. and McGuire, C.H., Use of Profile System for Scoring and Reporting Certifying Examinations in Orthopaedic Surgery. J. Med. Educ. 46 78, 1971.

Mager, R.F., Preparing Instructional Objectives. Fearon, 1967.

McGuire, C.H., The Oral Examination as an Assessment of Professional Competence. J. Med. Educ. 41 267, 1966.

Miller, G.E., The Orthopaedic Training Study. J.A.M.A. 206 601, 1968.

Nattress, L.W., Orthopaedic Certification Examinations. J. Med. Educ. 44 890, 1969.

Nattress, L.W., The Validity of Measurements of Clinical Competence. In Evaluation in Medical Education. Gilbert (Ed) Bulletin Commercial, Edmonton, 1971.

Popham, J.W. and Husek, T.R., Implications of Criterion Referenced Measurement. J. Educ. Meas., 6 1, 1969.

Popham, J.W., Selecting Objectives and Generating Test Items for Objective Based Tests. In Problems in C.R.M. Harris et al. (Eds), Centre for the Study of Eval. UCLA, 1974.

Shulman, L.S., Maatch, J.L. et al., Unpublished Grant Proposal, 1975.

Thorndike, R.L. (Ed), Educational Measurement (2nd edition), American Council

on Education, 1971.

Van Wart, A.D., A Problem Solving Oral Examination for Family Medicine. J. Med. Educ. 49 673, 1974.

5. MODEL FOR DEVELOPING A PROBLEM–BASED CRITERION REFERENCED TEST OF THE CLINICAL COMPETENCE OF MEDICAL STUDENTS

It should be emphasised that this model is incomplete and represents an approach developed by the author as a project commenced at the University of Southern California under the guidance of Professor Stephen Abrahamson. It has been found generally acceptable as a working model by several leading medical educators working in this field. The project is currently supported by a grant from the Education, Research & Development Committee (ERDC). In no way has it been sanctioned by the University of Adelaide.

The Development of the Project

The project development can be considered under several headings:

1. Literature Review
2. Site Visits
3. Definition of Clinical Competence
4. Statement of Objectives
5. The Problem Based Approach
6. Production of a Problem Blueprint
7. Selection of Test Methods
8. Production of an Examination Matrix

1. Literature Review

The outcome of this review is apparent in earlier chapters. This review was made possible only by the generosity of many educators in the field. Much of the data quoted has not been published in journal articles or books and is only available as U.S. government reports, in confidential reports or grant proposals and in limited circulation documents.

2. Site Visits

The author had the support of the University of Adelaide and the ERDC in visiting the following centres of medical education which enabled him to view at first hand most of the evaluation methods described and to develop skills in those selected as being of most potential value. Only those centres visited which are making a major contribution in the field of evaluating clinical competence are included.

Great Britain

Department of Medical Education, University of Dundee (Dr. R. Harden et al.)

Department of Audio-visual Communication, BMA, London

United States

Division of Research in Medical Education, University of Southern California (Prof. S. Abrahamson et al.)

Office of Health Sciences Education, University of California, San Francisco (Prof. E. Rosinski)

Office of Research in Medical Education, University of Washington (Dr. C. Dohner et al.)

Office for Medical Education Research and Development, Michigan State University (Drs. L. Shulman, N. Kagan, J. Maatch et al.)

Office for Educational Development, University of Illinois (Prof. C. McGuire et al.)

National Board of Medical Examiners (Dr. B. Andrew et al.)

Division of Research and Evaluation, Ohio State University (Dr. C.B. Melecca et al.)

Canada

Program for Educational Development, McMaster University (Dr. V. Neufeld et al.)

Department of Medicine, University of Manitoba (Dr. D. Ostrow)

Department of Educational Planning and Assessment, University of Calgary (Dr. L. Fisher et al.)

R.S. McLaughlin Examination and Research Centre, University of Alberta (Dr. D. Wilson et al.)

Health Science Office of Education, Queen's University (Dr. A. Bryans)

3. Definition of Clinical Competence

This area was reviewed in Chapter 2. However, some modification and development of this work has been undertaken to take into account the particular needs of evaluating competence in Australia's medical students. To try and clarify these particular needs, an attempt has been made to develop a statement of the goals of undergraduate medical education relating to clinical competence. In doing so, it was found necessary to identify as the end point, the production of a doctor capable of performing effectively as an intern (junior resident). It is recognised that academic staff will hope that the students develop many other capabilities and will have a wider perspective on medical practice. However, in terms of clinical competence, the intern level is the level at which he must function in practice and will be the primary motivating force at this stage in his career. It is certainly a major expectation of the clinical staff of the hospitals.

Statement of Goals (Job Description)

As a starting point in defining clinical competence, it is useful to attempt to define the broad goals of undergraduate training. This has been done in the form of a job description.

> "A recently graduated doctor is clinically competent. He is able to perform the clinical duties and administrative duties of a junior resident (intern) in a teaching hospital. When dealing with a patient, he is able to identify the patients' problems whether they are medical or psychosocial. He can write a reasoned assessment for each of these problems and can instigate appropriate investigations and treatment. In doing so, he demonstrates his appreciation of the fact that the patient is an individual.
>
> In managing the patients' problems he recognises his own limitations. He demonstrates this, firstly, by consulting his peers, registrars or consultants when necessary and, secondly, by utilising the skills and services of allied health professionals. He regards the evaluation of his performance as a normal and essential part of his practice. He is a continuous and self learner. He has developed skills that enable him to work effectively in a group or team. He is an effective teacher.
>
> The implications of his job extend beyond those relating to individual patient care. He knows that his practice must take into account its effect on the community he serves and for society as a whole."

It should be emphasised that this statement identifies the goals in terms of what the student will be expected to achieve.

Intern Job Structure

Having defined what it is aimed to achieve for the student, it is useful to look at the intern's job in more detail. A systems analysis type of flow diagram is a helpful way of visualising the job structure. In generating this diagram, thought has been given to its potential use for guiding the development of teaching strategies and evaluation procedures. It also includes the essential elements of current knowledge about the clinical problem solving process (See Fig. 1). It is, of course, an over-simplification.

Categorisation of Clinical Competence

It has been found to be most helpful to approach the detailed definition of clinical competence by using a categorisation of the process involved. The reasons for selecting process measures rather than outcome measures when evaluating medical students have been outlined previously. The classification developed by Hubbard et al. (1965) from the major critical incident study conducted for the NBME has been accepted with minor modification (Appendix 1). A tenth category – Administrative Responsibility – has been added, though this could have been included under "Responsibilities as a Physician". The sub-categories of this classification have been rewritten as behavioural statements which define the broad areas that constitute clinical competence (Appendix 4).

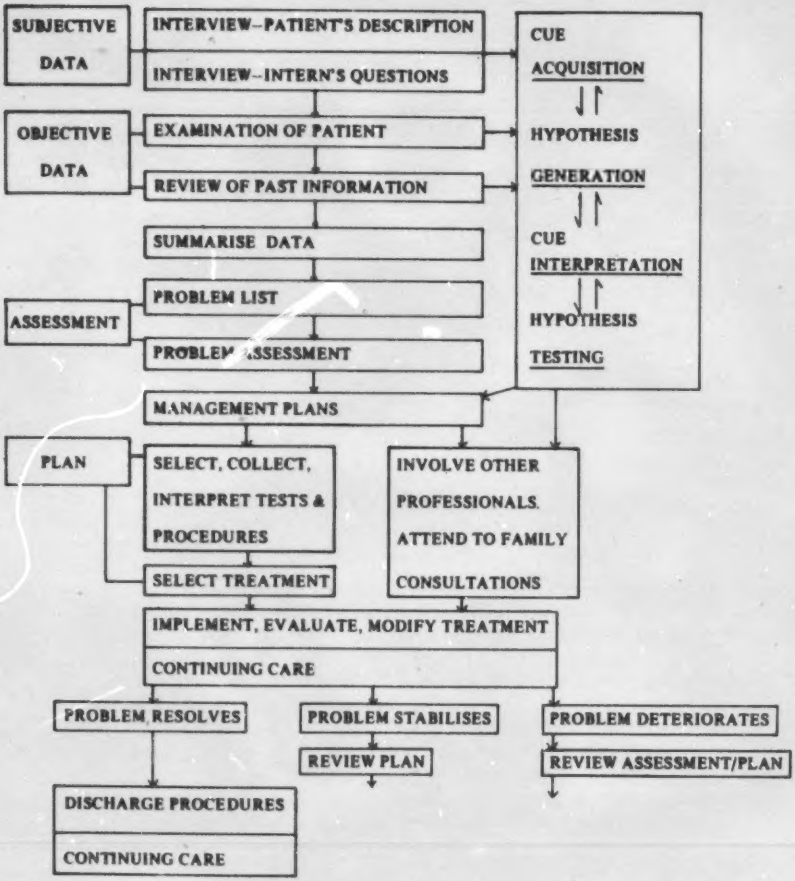
4. Statement of Objectives

It is necessary to expand this categorisation and write detailed statements which clearly identify the specific behaviours that the student must demonstrate. The work in this area is incomplete but the preliminary documents relating to categories I, II and VII have been prepared (Appendix 5). It may well be that these documents reflect the bias of a Physician (internist) and the final development of behavioural objectives should satisfy doctors from all specialties including general practitioners and perhaps even representatives of other health professions and the community. It is essential that the clinically competent final year student be seen as a "whole" rather than as a conglomerate of embryo specialists or as a budding general practitioner. The development of these more specific and specialised roles is properly one of postgraduate training commencing after the intern year.

Finally, it should be noted that no effort has been made, as yet, to set exact standards or criteria of behaviour. Such details would require the setting up of one or more task forces of clinical experts from all disciplines or the use of a consensus questionnaire approach like the Delphi technique (Delbecq et al., 1975).

FIGURE 1

CLINICAL RESPONSIBILITIES OF AN INTERN



5. The Problem Based Approach

The traditional approach to curriculum planning has involved identifying subject areas and leaving departments to identify the subject matter. Recent trends have been towards body system or topic teaching with varyingly successful efforts at integrating departmental teaching. The major emphasis is still on content knowledge. This is the current situation at the University of Adelaide. An alternative approach, developed to the greatest extent at McMaster University and Michigan State University, bases the curriculum on developing skills at solving problems. At McMaster it is assumed that if an adequate number of problems is covered the, inevitably, relevant content knowledge will be acquired and that this knowledge will be more easily remembered (Neufeld & Barrows, 1974). No one can argue the validity of such an approach but the total McMaster philosophy and approach would be difficult to institute in an established and more traditional medical school. However, solving patient problems is what an intern must do, so any evaluation of clinical competence which does not include problem solving tests could hardly be considered valid. A problem based approach has therefore been selected as the basis of the examination model.

At the end of 1974, the Terminal Learning Objectives for the M.D. Programme at the University of Manitoba were obtained from Dr. David Ostrow. These provided an approach to objective and problem definition that intuitively seemed to make sense. The objectives defined for the Department of Medicine were sent, unofficially, to members of the teaching faculty with particular responsibilities in each major specialty area. They were asked to comment on the approach and make any alterations they saw fit. There was almost unanimous approval for the approach and many helpful suggestions and minor alterations were made. The document circulated is reproduced, without modification, in Appendix 6. It should be noted that the cognitive (knowledge) area is composed of problems and tests and that these are followed by identified skills and attitudes that they feel the student should achieve.

Accepting the problem based approach to evaluation, it is now necessary to develop a method for producing an outline or definition of the specific knowledge and skills which are required for the doctor to successfully handle each problem. The definition of objectives has often been the stumbling block in curriculum development. At this stage, one has to face the fact that a much larger section of the clinical, and even preclinical, faculty will need to be involved. They are unlikely to give this activity a high priority and time constraints will limit their involvement. Implementation of innovative change has often faltered or totally foundered because the burden of the change was too great or the concepts were too complex to be readily understood. In an attempt to avoid these pitfalls a procedure has been developed which should allow a subject specialist to more easily identify and define the key areas of knowledge or skills which would be required to solve each problem. This procedure leads to the production of a "problem blueprint".

6. Production of a Problem Blueprint

A problem is something the patient has. The patient may be aware of it himself (e.g. chest pain, headache, shortness of breath) or it may become apparent only during the doctor–patient contact (e.g. hypertension, cardiac murmer, an abnormal blood test). Whatever the problem, it has widely differing implications for the doctor in terms of the skills he must master to solve the problem. One problem may require a skilled interviewing technique but little therapeutic skill, another may require instant emergency action and yet another may require a detailed understanding of therapeutics but little diagnostic skill. Clearly, each problem is an entity and this may explain why there is only a low correlation between problem solving tests designed around different problems.

This difficulty has been tackled by developing the concept of a "problem blueprint". An individual problem is considered from the standpoint of each of the 10 major categories of clinical competence previously defined (Appendix 1). The detailed objectives under each category are used as a checklist and relevant features that the subject expert feels may be required to solve the problem are tabulated. This procedure can be followed independently by several subject experts to check congruence. The same can be done by clinicians in different specialties seeing the problem from different or more general perspectives and also by a variety of non-physicians who may have much to contribute, e.g. social workers, nursing staff, patients.

An example of such a blueprint is seen in Appendix 7. The problem taken was "chest pain – cardiac" and the perspective is that of a General Physician (Internist) in a full-time academic position. The blueprint would be different if prepared by a cardiologist, a coronary care nurse, a patient who had previously had a myocardial infarction and so on. It would be the final responsibility of the Faculty to determine what were the relevant areas for evaluating a medical student and to what standard. The preparation of such a blueprint is simple and quick and would not an undue burden even if one were asked to do the same for several problems. It provides a ready reference index to the key areas of content knowledge, technical skills and attitudes that are required. If thought worthwhile, a behavioural objective could be written for each point but though this might please the purist, it may not produce much pay-off for the Faculty.

7. Selection of Test Methods

There are two aspects to selecting test methods that need to be considered. Ideally it should be possible to match up a single test method to each category of clinical competence. In practice, the specificity of the tests currently available does not allow this to be done with such precision but some techniques are clearly more appropriate to evaluating one category rather than another. The second consideration is one including practicability, convenience, cost and availability of personnel. A compromise will always be required but every attempt should be made to select the best method available. This requires that the faculty staff responsible for setting examinations have an adequate knowledge and understanding of test methodology and preferably have access to a measurement and evaluation expert.

A chart showing some of the techniques currently being used is included in Appendix 8. Although incomplete, it does give some idea of how test methods might be matched against each category of competence and demonstrates the overlap of techniques between categories. Some of the techniques would not be feasible to include in a traditional final examination. This does not mean they should be discarded but means that the structure of final examination might need changing. It might be inappropriate to attempt to assess some areas of competence in a single examination sitting. Some may need to be assessed over a more prolonged period of time. Achieving competence in these areas would then be a pre-requisite to sitting the final examination.

8. Production of an Examination

It is obvious that no examination can cover all aspects of the subject. Some sampling procedure is inherent in all test situations and it is hoped that valid generalisations about ability in the whole field can be made from the results. This is an area requiring more thought than has been the custom. Students will quickly identify any pattern of testing that is biased or non-random and will adjust their learning behaviour and examination-taking techniques accordingly. A mechanism must be incorporated in the model to overcome such difficulties. If we assume, once again, that the problem based approach is in effect, all students will have a list of problems and these will not all fit neatly into a specialty based format. Well ahead of the examination, say at least six months, a selection of problems would be chosen randomly. A non-departmental examination committee could then set about producing problem based items. A selection, random if possible, of key items from each category of the problem blueprint would be made and the appropriate test methods selected. This would lead to a multiple method examination format similar to that being used by the American Board of Orthopaedic Surgeons (ABOS) and the Canadian College of Family Practice (CCFP). However, in both these cases, many of the test items will be predetermined and not randomly selected. Finally, a profile scoring system would be developed which would differ quite radically from the ABOS and CCFP profiles. These bodies attempt to evaluate a series of general abilities, such as recall, interpretation, problem solving and attitudes, while the proposed model presents a profile score which would be concerned with more specific competence categories.

It is appreciated that an approach like this implies a considerable amount of work for the examination committee but it is expected that it would have a major benefit on the attitudes of students and their terminal performance and could have a tremendous impact on the whole curriculum. An independent examination would clearly identify areas where competence was not achieved and feedback could be channelled back to the appropriate curriculum or management committee.

The concepts of producing such an examination matrix could best be illustrated by examples referable to the problem blueprint for chest pain (Appendix 7). The following tests might be used:

a. Multiple choice questions could be sought to test knowledge about trinitrin and betablockers, the treatment of arrhythmias; identification of risk factors, etc.

b. Short PMP's could be written to test problem assessment, use of tests and procedures, other aspects of treatment, etc.

c. Slides of X-rays, ECG's, monitor tracings could be projected for interpretation.

d. Real or simulated patients could be available to evaluate history and examination skills and to provide a data base for a case write up.

e. Audiotapes of ausculatory findings could be combined with a PMP or a series of objective items.

f. A simulation of a cardiac arrest could be arranged or a resuscitation mannequin used to observe life support skills.

g. Structured orals could cover areas dealing with continuing care, effective patient education, assessing awareness of recent trends, etc.

With suitable administrative arrangements, the overall time spent, at least by examiners, should not be a great deal more than previously. A certain amount of examiner training would be required to improve reliability. The objective structured examination described by Harden et al. (1975) using multiple ward stations demonstrates the feasibility of this examination format. The outcome of such an examination would be a separate analysis of performance in each area of competence over a variety of problems. A profile score for each student would be available to the examiners and, hopefully, would be made available to the students. It would, of course, be ideal if a similar examination could be conducted several weeks or months before the final or certification examination to produce a diagnostic profile primarily for student feedback. Remedial work could then be undertaken prior to the finals.

REFERENCES

Delbecq, A.L., Van de Ven, A.H. and Gustafson, D.H., Group Techniques for Program Planning. A Guide to Nominal Group and Delphi Processes. Scott, Foresman & Co., 1975.

Harden, R.McG., Stevenson, M., Downie, W.W. and Wilson, G.M., Assessment of Clinical Competence using an Objective Structured Examination. Brit. Med. J. 1 447, 1975.

Hubbard, J.P., Levit, E.J., Schumacher, C.F. and Schnabel, T.G., An Objective Evaluation of Clinical Competence. New Engl. J. Med. 272 1321, 1965.

Neufeld, V.R. and Barrows, H.S. The McMaster Philosophy: An Approach to Medical Education. J. Med. Educ. 49 1040, 1974.

APPENDIX 1

CATEGORIES OF CLINICAL PERFORMANCE

1. **History:**

   Obtaining information from patient.
   Obtaining information from other sources.
   Using judgment.

2. **Physical Examination:**

   Performing thorough physical examination.
   Noting manifest signs.
   Using appropriate technique.

3. **Tests and Procedures:**

   Utilising appropriate tests and procedures.
   Applying test methods correctly.
   Modifying tests to meet the patient's needs.
   Interpreting test results.

4. **Diagnostic Acumen:**

   Recognising causes.
   Exploring condition thoroughly.
   Arriving at a reasonable differential diagnosis.

5. **Treatment:**

   Instituting the appropriate type of treatment.
   Deciding on immediacy of the needs for therapy.
   Judging the appropriate extent of treatment.

6. **Judgment and Skill in Implementing Care:**

Making necessary preparations.
Using correct methods and procedures.
Performing manual techniques properly.
Adapting method to special procedures.

7. **Continuing Care:**

Following the patient's progress.
Modifying treatment.
Planning effective follow-up care.

8. **Physician–Patient Relation:**

Establishing rapport with the patient.
Relieving tensions.
Improving patient co-operation.

9. **Responsibilities as Physician:**

For the welfare of the patient.
For the hospital.
For the health of the community.
For the medical profession.

(Hubbard et al., 1965)

APPENDIX 2

I. COLLECTING AND INTERPRETING CLINICAL INFORMATION

A. Obtaining Information from the Patient.

1. Obtaining adequate medical history of patient and/or family.
2. Obtaining relevant social and emotional history.

B. Checking Multiple Sources for History.

1. Questioning family members.
2. Consulting other sources.

C. Obtaining Information by Physical Examination.

1. Doing indicated physical examination.
2. Performing examination correctly.

D. Obtaining Information by Special Tests.

1. Obtaining indicated test data.
2. Applying and interpreting test properly.

E. Approaching Diagnosis Objectively.

1. Double checking stated or referral diagnosis.
2. Avoiding prejudgments.
3. Giving appropriate emphasis to emotional and physical illness.

F. Recognising Condition.

1. Recognising specific medical problem.
2. Recognising emotional symptoms.
3. Recognising emotional basis of physical complaint.

II. INSTITUTING AND MAINTAINING COMPREHENSIVE CARE

A. Administering Therapy.

1. Suitably treating condition.
2. Adapting treatment to special needs.
3. Acting with composure and effectiveness in emergencies.

B. Treating More Than Immediate Health Problem.

1. Modifying patient's physical or social environment.
2. Dealing with problems of patient and family.
3. Instituting preventive measures.

C. Working with Health Care Team.

1. Making appropriate use of health care team.
2. Integrating efforts of team members.

D. Providing Follow-Up Care.

1. Providing for rehabilitative care.
2. Planning home or self-care.
3. Instructing patient and family in home care.

III. PHYSICIAN–PATIENT RELATIONSHIP

A. Establishing and Maintaining Suitable Relationship.

1. Overcoming communication problems.
2. Acting in discreet, tactful, and dignified manner.
3. Encouraging patient expression.
4. Exploring patient's statements and ideas.

B. Showing Consideration for Patient and Family.

1. Avoiding needless pain, discomfort, or embarrassment.
2. Avoiding needless alarm.
3. Avoiding needless inconvenience.

C. Relieving Anxiety of Patient and Family.

1. Reassuring, supporting or calming.
2. Giving adequate explanation of condition, treatment and prognosis.

D. Insuring Needed Care.

1. Persuading patient to undertake needed care.
2. Eliciting reasons for failure of care.
3. Insuring needed medication or hospitalisation.
4. Making health care accessible.

IV. RESPONSIBILITIES OF A PHYSICIAN

A. Accepting Responsibility for Welfare of Patients.

1. Heeding the call for help.
2. Devoting necessary time and effort.
3. Meeting commitments.

B. Maintaining Objectivity.

1. Avoiding bias.
2. Keeping appropriate psychological distance.
3. Handling own emotions.

C. Recognising Professional Capabilities and Limitations.

1. Doing only what experience permits.
2. Asking for help, advice or consultation.
3. Developing own medical knowledge and skills.

D. Accepting Responsibility for Community Health.

1. Recognising health needs of other family members.
2. Examining and treating other family members.
3. Preventing spread of infection among hospital patients.
4. Preventing illness in patient's family and others in community.

(Sanazaro & Bates, 1968)

APPENDIX 3

CRITERIA FOR CLASSIFYING M.D. ACTIONS

1. GENERAL CARE

Professional manner	Physician behaviour/attitudes in dealing with patients; establishing or maintaining rapport, or failure to do so.
Physician availability	Making house calls, responding to calls from nurse, hospital, patients; seeing patient personally; attending patient; willingness to see patient; providing for coverage.
Use of health team	Co-ordinating services of other physicians, nurses, auxiliary workers; promoting, facilitating communication among professionals.
Use of facilities	Hospitalising or transferring patient; use of specialised equipment.
Follow-Up	Arranging, assuring, scheduling follow-up care; making effort to do same; attitude toward providing such care,
Professional responsibility	Matters of ethics, interprofessional relationships and attitudes; continuing education.
Consultation	Use of consultation for diagnosis and/or treatment; use of advice offered by consultants.
Other	M:D. actions which clearly relate to General Care but are not classifiable under Evaluation or Treatment.

2. EVALUATION: GENERAL

Review of problem	Re-evaluation, re-examination, re-investigation; review of records; discussing with prior physicians.
Review of treatment	Reassessment of regimen.
Arriving at diagnosis	Making, considering diagnosis or differential diagnosis; awareness or recognition of causes, conditions, diagnoses.
Arriving at plan of treatment	Deciding upon a plan of treatment, for example, to use drugs, surgery, dialysis, etc.
Other	Evaluation of problem or situation by unspecified methods; diagnostic work-up; exploring problem (if not clearly psychological); "examination"; investigation of problem, situation, reaction, cause.

3. EVALUATION: SPECIFIC

Psychological perception	Recognition by M.D. of patients' special psychological or emotional needs, attitudes, beliefs, 'unspoken' complaints.
History	All actions that pertain to eliciting, recording, verifying, interpreting history; use of history, attitudes toward history-taking.
Physical examination	Physical examination rather than 'examination' undefined; acts of noting, discovering, missing, finding, performing; attitudes toward; techniques.
Use of instruments	Use of any instrument in examining patient that extends the native senses: ophthalmoscope, otoscope, stethoscope, anuscope, proctoscope.
Laboratory	Refers both to use of laboratory, as in ordering tests, and use or interpretation of results; reliance; selectivity.

X-ray	Refers both to use of radiological techniques and to interpretation or use of interpretations of X-rays.
EKG	Same as for X-ray.
Diagnostic procedures	Biopsy, Papanicolaou smears, lumbar puncture.
Other	

4. **TREATMENT: GENERAL**

Revising treatment	Actual revision of attitude or habitual approach.
Other	

5. **TREATMENT: SPECIFIC**

Community resources	Use of special agencies, community health facilities, family services, child guidance, visiting nurse association.
Patient education	Instructing, informing, educating, explaining, preparing patients. Primary purpose is information exchange, in distinction to 'psychological support'. Goal is increased patient knowledge, understanding, self-care, compliance.
Psychological support	Working directly with patient or family to reassure, support, alleviate concern, counsel; expression of interest in patient, family. Goal is improved emotional state, greater insight, security.
Physical modality	Cardiac resuscitation; hemodialysis; artificial respiration; peritoneal dialysis, physical therapy, massage.
Procedure	Non-surgical procedure used in treatment; removal of foreign body from trachea or body cavity; phlebotomy; local injection.
Radiation–electrical	Use of X-ray therapy in any form; other wave therapy; cardiac defibrillation; ultrasonic or heat therapy.

Surgical treatment	The immediate preoperative, intraoperative and postoperative treatment, including technique; judgment; recommending; refusing; attitude toward.
Drugs, biologicals, electrolytes, fluids	Administering, prescribing, ordering, recommending drugs, biologicals, fluids, electrolytes; knowledge of; dose; monitoring; side effects.
Diet	Diet as the modality of treatment.
Prescription of activity	Specific advice to increase, limit, modify, regulate activity. Includes prescription of bed rest, rest periods, use of commode.
Other	Any other specific modality of treatment not included elsewhere.

(Sanazaro & Williamson, 1968)

APPENDIX 4

DEFINITION OF CLINICAL COMPETENCE

MAJOR CATEGORIES

I.	Obtaining the Data Base	The graduating student will be able to obtain both the subjective and objective information that contribute to a comprehensive data base. This will be obtained and recorded in the case record in a systematic and problem-oriented manner. The student will use the history from the patient, information from other sources, the physical examination findings, and data obtained from previous medical investigations.
II.	Problem Assessment Diagnosic Acumen	The graduating student will be able to accurately assess and write down the assessment of problems identified from the data base. He will recognise causes of the problems and, after exploring possible alternatives, will arrive at a reasonable differential diagnosis.
III.	Tests and Procedures	The graduating student will have a knowledge of routine and special tests and procedures currently available in his major teaching hospitals. He will be able to utilise and apply the test methods correctly to identified problems and to modify them to meet the patients' needs. He will be able to interpret all routine tests but not necessarily all special tests.
IV.	Treatment	The graduating student will institute the appropriate type of treatment and will be able to decide on the immediacy of the needs for therapy. He will be able to judge the appropriate extent of treatment.

V.	Judgment and Skill in Implementing Care	The graduating student will show judgment and skill in implementing investigation, management and treatment plans. He will be able to make any necessary preparations, to use correct methods and procedures, and to properly perform manual techniques.
VI.	Continuing Care	The graduating student will be able to take responsibility for the continuing care of patients. He will follow the patient's progress and when appropriate will modify his assessment, plan of management and treatment. He will be able to plan effective follow-up.
VII.	Doctor–Patient Relationship	The graduating student will be able to establish rapport with the patient and relieve tensions where present. He will be able to improve patient co-operation.
VIII.	Responsibilities as a Doctor	The graduating student will be prepared to undertake the responsibilities of a doctor. These include those relating to the welfare of the patient, the hospital or practice, the community and society, and the medical profession.
IX.	Continuing Education and Evaluation	The graduating student will regard the evaluation of his clinical performance as a normal and essential part of his practice. He will actively engage in appropriate and effective educational activities.
X.	Administrative Responsibilities	The graduating student will be familiar with and accept the administrative duties of an intern (junior resident).

APPENDIX 5

STATEMENT OF BEHAVIOURAL OBJECTIVES

CATEGORY I. OBTAINING THE DATA BASE

A. Subjective Data From the Patient.

1. Establishes rapport.
2. Conducts the clinical interview, as far as possible, in a non-directive manner.
3. Identifies the major problem of concern to the patient.
4. Identifies other problems having present or potential relevance.
5. Mentally classifies the reliability of the patients' information.
6. Identifies the information relevant to the major and other problems.
7. Obtains a complete screening history of areas where no obvious problems have been identified.
8. Obtains a history which gives an adequate profile of the patients' life-style, socio-economic status, ethnic background, mental status, family history and past medical and drug history.
9. Records the subjective data base accurately, precisely, comprehensively, comprehendably and legibly.
10. Records the data in a problem-oriented manner identifying the sequence of development of symptoms, their characteristics and interrelationships.
11. Checks his perceptions of the data with the patient.

MODIFYING CIRCUMSTANCES

a. Psychiatric patients.
b. Children.
c. Adolescents.
d. Emergency situations.
e. Short consultations.
f. Follow-up patients.
g. Language problems.
h. Speech problems.
i. Mental deficiency.
j. Comatose or semi-comatose patient.

B. **Subjective Data From Other Sources.**

1. Identifies areas in the patients history which need confirmation or expansion from an additional sources.
2. Identifies the most appropriate source.
3. Conducts family interviews when appropriate.
4. Records this data in the case record.

C. **Objective Data From the Patient.**

1. Continues rapport.
2. Conducts the physical examination efficiently and minimises physical and mental discomfort to the patient.
3. Performs the examination on a regional basis.
4. Performs a detailed examination of the organ system or body area which relates to the identified problems.
5. Completes a screening examination of the other organ systems or body areas.
6. Identifies all abnormal physical findings but seeks confirmation of doubtful findings.
7. Identifies relevant negative findings.
8. Recognises new problems and takes additional history related to these problems.
9. Able to use the standard diagnostic instruments to identify normal and abnormal findings.
10. Performs a mental status examination when appropriate.
11. Identifies information relevant to the major and other problems.
12. Records the objective data in a systematic manner in the case record.

MODIFYING CIRCUMSTANCES

a. Psychiatric patients.
b. Children.
c. The opposite sex.
d. Emergency situations.
e. Short consultations.
f. Follow-up patients.
g. Deformed, debilitated or immobilised patients.
h. Comatose or semi-comatose patients.
i. Language or receptive aphasia problems.

D. **Objective Data Obtained From Other Sources.**

1. Seeks appropriate data from other sources.
2. Validates any data obtained from other sources.

E. The Case Record.

1. Appreciates the importance of the case record.
2. Writes legibly.
3. Records all data in the format appropriate to the hospitals record system.
4. Writes a succint summary of the patients' presenting problems and data.
5. Constructs a complete problem list indicating active and inactive problems.

CATEGORY II. PROBLEM ASSESSMENT AND DIAGNOSTIC ACUMEN

A. Identification of Problems.

1. Identifies all the medical and psycho-social problems which can be obtained from the data base.
2. Writes down a complete but concise master problem list in the problem oriented medical record, using the prescribed format.
3. Distinguishes the essential features or cues in the data base which are relevant to each problem.
4. Write these features and cues under Subjective(S) and Objective(O) headings for each defined problem in the medical record.
5. Recognises that the patient may have problems that he has as yet not identified.

B. Assessment and Diagnosis of Problems

1. Understands the principal steps of the problem-solving process and how the problem oriented medical record can be effectively used to help in this process.
2. Formulates hypotheses which are applicable to the diagnosis of the patients' problems. He starts to do this from the beginning of the clinical interview.
3. Constantly attempts to verify the hypotheses during the initial data base collection.
4. Identifies the diagnostic cues that have discriminatory value.
5. Has a knowledge of the concepts, rules or principles relevant to the problems he is likely to encounter in his practice as an intern and how to seek these out when he does not have the required knowledge.
6. Develops a rank order of diagnostic possibilities for each problem based on a reasoned estimate of probabilities and including a consideration of potential seriousness and treatability of the considered diagnosis.

7. Knows when his assessment is incomplete or inadequate and takes appropriate steps to rectify the situation.
8. Remains aware that the hypotheses or diagnoses may have to be changed.
9. Writes an assessment that is readily perceived by all members of the patients' health team.

CATEGORY III. DOCTOR–PATIENT RELATIONSHIP

A. Establishing Rapport

1. Greets patient, introduces himself and defines his role.
2. Transcends age, social and cultural barriers.
3. Inspires a feeling of trust, care and compassion.
4. Understands the social and cultural context of the illness for the patient.
5. Develops an atmosphere conducive to pontaneous behaviour.

B. Relieving Tension

1. Understands the emotional and behavioural responses to illness.
2. Not upset by emotional and behavioural responses directed personally.
3. Explores these responses with the patient.
4. Shows by verbal and non-verbal means that he recognises and is concerned by the patients anxiety and discomfort.
5. Attempts to provide a positive feeling of well-being at a realistic level.
6. Copes with difficult situations in the doctor–patient relationship.
7. Is a skilful listener.

C. Improving Co-operation and Compliance

1. Able to communicate, verbally and non-verbally, with the patient.
2. Willing to continue the relationship when indicated or to ensure the transfer of care to others.
3. Works towards a maximum of mutual participation in management.
4. Tailors the management plan to suit the individual patient.
5. Checks the patients understanding of the treatment plan.

APPENDIX 6

TERMINAL LEARNING OBJECTIVES
DEPARTMENT OF MEDICINE
UNIVERSITY OF MANITOBA

A. Cognitive Area:

Given a patient presenting with any of the problems listed below, the student should be able to:

1. OBTAIN a comprehensive problem-oriented history.

2. PERFORM a complete problem-oriented physical examination.

3. DEMONSTRATE competent problem-solving skills.

   Competent problem-solving skills require a knowledge of and movement from general pathophysiologic derangements to specific disease entities. Students should be able to AT LEAST DISCUSS THE BASIC PATHOPHYSIOLOGIC DERANGEMENTS IN THE PROBLEMS.

   A simplified outline of problem-solving skills might be:

collection of information from patient	DATA COLLECTION
compare with known physiologic variations	
if abnormality present	PROBLEM SENSING
hypothesis generation	
hypothesis testing	SOLUTION TESTING
treatment / revision of hypothesis	

4. DEMONSTRATE where they would define the limit of their own competence in a problem and when they would seek advice and consultation in the problem management.

5. PRODUCE a plan of management including:

   a. ranking problems in terms of priority along with a differential diagnosis of each problem

   b. appropriate diagnostic procedures to hypothesis test

   c. most appropriate currently prescribed medications, and their pharmacology, indications, dosages, side effects, toxic symptoms and signs

   d. a plan for patient education and information transmission to relatives

   e. consultation, where appropriate, of available health resources, such as professionals in the hospital, clinic and community who may have cared for the patient and who contribute to the continuing care of the patient

   f. a problem oriented record.

6. ANALYZE the factors in the environment, internal and external, that predispose to disease and PROPOSE a plan for eradicating these and a standard programme for preventative medicine.

Problems:

1. Anemia
2. Polycythemia
3. Bleeding tendency
4. Lymphadenopathy and Splenomegaly
5. Alterations in leukocytes

6. Sore Throat
7. Cough, Sputum, Hemoptysis
8. Shortness of Breath
9. Abnormalities of Chest X-ray – especially coin lesions, parenchymal infiltrates, cavities

10. Wheezing
11. Cyanosis
12. Chest Pain
13. Palpitations and abnormalities of rhythm
14. Cardiac and other murmurs
15. Hypertension
16. Hypotension

17. Shock
18. Sudden death

19. Abdominal Pain – acute, chronic
20. Anorexia, Nausea, Vomiting
21. Dysphagia
22. Indigestion
23. Hematemesis and Melena
24. Constipation and Diarrhea
25. Obesity
26. Abdominal Mass(es)
27. Hepatomegaly
28. Weight Loss
29. Ascites and swollen abdomen
30. Jaundice

31. Polyuria, oliguria, nocturia
32. Dysuria
33. Hematuria
34. Proteinuria
35. Glycosuria

36. Infertility
37. Disturbances in sexual function

38. Painful extremities
39. Swollen extremities
40. Pain in the back and neck
41. Ill-defined musculoskeletal disorders

42. Acute septic illness
43. Fever of unknown origin
44. Recurrent infections

45. Allergic reactions

46. Itching
47. Alopecia and hirsuitism
48. Pigmentation
49. Photosensitivity
50. Rashes/skin eruptions
51. Nail bed changes

52. The Red Eye
53. Changes in Vision

54. Mass in the Neck
55. Changes in Hearing

56. Headache

57. Sleep abnormalities
58. Convulsions
59. Disturbances in Sensation
60. Syncope
61. Giddiness, vertigo
62. Derangement of consciousness, coma
63. Delerium and confusional states
64. Derangements of intellect and behaviour
65. Lassitude, asthenia and generalised weakness
66. Disturbances of Gait
67. Disturbances of Speech
68. Disturbances of Balance
69. Localised weakness

70. Nervousness
71. Anxiety
72. Depression

73. Acute and Chronic Intoxications
74. The terminally ill patient

The student should be able to INTERPRET the following:

1. Common EKG abnormalities
2. Blood gas reports
3. Routine pulmonary function reports
4. Common electrolytic patterns for blood and urine
5. Acid base status reports
6. Bone marrow interpretation

The student should display an ability to analyse all factors of health and disease in the community environment.

B. Skills:

Upon graduation, the student should be able to PERFORM a complete examination of a patient including the following:

1. Pelvic examination and use of a speculum
2. Indirect laryngoscopy
3. Opthalmoscopy/otoscopy
4. Measurement of intraocular pressure with a tonometer
5. Sigmoidoscopy/proctoscopy

The student should be able to PERFORM and INTERPRET:

1. Sputum collection and Gram stain evaluation
2. Stool examination for occult blood
3. Preparation and observation of blood smears

4. Urinalysis (examination of stained and unstained sediment and analysis with a dip stick)
5. Lumbar puncture with cell count and Gram stain
6. Hemoglobin estimation with a Spencer Hemoglobinometer
7. Blood smear and WBC and differential
8. Collection of specimen for investigation of V.D. in either sex

The student should be able to PERFORM and DISCUSS their indications and associated hazards:

1. Nasogastric intubation
2. Endotracheal intubation
3. Throat culture and plating
4. Catheterisation of male and female bladder
5. Venipuncture/I.V. infusion
6. Arterial puncture
7. Closed chest massage/cardiopulmonary resuscitation
8. Bone marrow aspiration
9. Injections – intradermal; subcutaneous; intramuscular
10. Thoracentesis
11. Central venous pressure measurement
12. Nasal packing

The student should have OBSERVED and should be able to describe the following procedures, their indications and their hazards:

1. Pericardiocentesis
2. Tracheostomy
3. Needle tracheostomy
4. Liver biopsy
5. Kidney biopsy
6. Pleural biopsy
7. Sigmoidoscopic biopsy of a lesion
8. Chest tube for a pneumothorax

C. Attitudes:

Given a patient with any of the problems listed in A, a graduating student should demonstrate actions that:

1. a physician's own continuing education is an important personal responsibility

2. reassure the patient that he has accepted the responsibility for management of the problem

3. diagnosis precedes definitive therapy except in certain specific cases where the risks of the diagnostic process are greater than the risks of incorrect therapy

4. show that the effectiveness of a physician's own medical care in influenced by his approachability and sense of concern for his patient's welfare

5. he informs patients of their diagnosis and general prognosis at a point in time when their adjustment permits and he informs them of reasons for proposed diagnostic studies

6. the provision of health care is the responsibility of many health professionals of which the physician is only one and that he respects, consults/and co-operates with other members of the team in a realistic way

7. willingness to be objectively criticised and to criticise objectively and constructively his peers.

APPENDIX 7

PROBLEM BLUEPRINT
CHEST PAIN – CARDIAC

MAJOR CATEGORIES

Obtaining the Data Base	Cardio-respiratory history Cardio-respiratory examination Auscultation	
Problem Assessment/ Diagnostic Acumen	Knowledge of diagnostically critical symptoms and signs Generation of a diagnostic hypotheses in classical and atypical cases Definition of a problem list and weighting them on the basis of the current data base	
Tests and Procedures	ECG CXR Enzyme studies Antiography/Scanning	Monitoring Cardiac catheterisation Emergency life support Identification of risk factors
Treatment	Patient education Trinitrin Beta-blockers Pain relief Surgery	Complications Cardiac arrest Arrhythmias Heart failure Risk factors
Judgment and Skill in Implementing Care	Overall appropriateness of the assessment in an individual patient	
Continuing Care	Recognition and management of early and late medical and psycho-social complications Rehabilitation Effective patient education – including risk factors Co-ordination of care	

Doctor–Patient Relationship	Effective patient education Avoidance of dependancy and 'cardiac neurosis'
Responsibilities as a Doctor	Community education Assuring continuity of care Use of resources Education of nursing and paramedical staff
Continuing Education and Evaluation	Auditing of records and clinical performance Use of reference material and cardiologist consultation Awareness of recent trends
Administrative Responsibilities	Completes the case record and summary Knowledge of hospital administrative procedures

APPENDIX 8

ASSESSMENT OF CLINICAL COMPETENCE – TEST SELECTION

MAJOR CATEGORIES

Obtaining the Data Base	Direct observation of real or simulated patients using rating forms PMP's (paper, computer) Observation of selected videotapes or films Record audit
Problem Assessment Diagnostic Acumen	PMP's — Problem solving structured orals Objective tests — Simulations Record Audit Observation of videotapes, stimulus films, slides, X-rays, etc.
Tests and Procedures	PMP's — Record audit Objective tests Simulations Problem solving orals
Treatment	PMP's Objective tests Simulations Problem solving orals
Judgment and Skill in Implementing Care	PMP's — Psychological tests Direct observation Record audit Role playing orals
Continuing Care	PMP Record audit – outpatients Role playing orals

Doctor–Patient Relationship	Direct observation Role playing interviews Psychological tests Interaction process analysis	Patient ratings
Responsibilities As a Doctor	Peer rating Record audit Direct observation of ward performance Patient/Nurse ratings	Project assessment
Continuing Education and Evaluation	Direct observation of performance during audits Participation record in educational programmes	
Administrative Responsibilities	Objective tests Record audit	